



ARTICLE

DOI: 10.1038/s41467-018-07089-x

OPEN

An interpretable approach for social network formation among heterogeneous agents

Yuan Yuan ¹, Ahmad Alabdulkareem² & Alex ‘Sandy’ Pentland ³

Understanding the mechanisms of network formation is central in social network analysis. Network formation has been studied in many research fields with their different focuses; for example, network embedding algorithms in machine learning literature consider broad heterogeneity among agents while the social sciences emphasize the interpretability of link formation mechanisms. Here we propose a social network formation model that integrates methods in multiple disciplines and retain both heterogeneity and interpretability. We represent each agent by an “endowment vector” that encapsulates their features and use game-theoretical methods to model the utility of link formation. After applying machine learning methods, we further analyze our model by examining micro- and macro- level properties of social networks as most agent-based models do. Our work contributes to the literature on network formation by combining the methods in game theory, agent-based modeling, machine learning, and computational sociology.

¹Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ²Center for Complex Engineering Systems, King Abdulaziz City for Science and Technology and Massachusetts Institute of Technology, Riyadh 12354, Saudi Arabia. ³Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. Correspondence and requests for materials should be addressed to A.S.P. (email: pentland@mit.edu)

Social networks have attracted increasing attention from both physical and social scientists^{1–4}. Social networks are essential elements in societies, serving as channels for exchanging various benefits, such as innovation, information, and social support^{5–8}. Moreover, research in social networks helps explain macro-level social phenomena, such as social polarization⁹ and social contagion^{10,11}. An understanding of social networks has significant implications, such as improving social welfare and political participation^{12,13}.

Previous work on modeling social network formation has typically employed game theory or agent-based modeling^{14–20}. These studies typically propose simple and tractable micro-level rules for link formation mechanisms and show that these rules have implications for known macro-level properties. Several studies in statistics and econometrics have also used game theory to model empirical networks^{21–23}, but they typically have been focused on estimating and identifying the effects of interest, such as racial segregation. To date, these models have not been capable of accounting for the effects of broad heterogeneity among individuals; therefore, they lack predictive power for link formation in complex, real-world networks.

Studies on network embedding techniques^{24–27} could partially fill this gap in the network formation literature because these techniques consider node heterogeneity and show predictability of both link formation and individual characteristics. Network embedding techniques are aimed at representing each node with a fixed-length vector learned from social network data. The agents in a network may be so diverse that representing all their characteristics would require very high dimensionality for these vectors. The philosophy of network embedding is aimed at reducing the dimensionality by mapping all the characteristics of agents onto a low-dimensional latent space. Each dimension in the latent space, therefore, typically does not correspond to a concrete attribute of the agents. The latent space representation of nodes on a network provides considerable potential for measuring heterogeneity among agents. However, because network embedding methods are designed for data representation and compression rather than for explaining network formation, they do not attempt to capture micro, inter-agent effects such as social status or macro effects such as social segregation; thus, they do not provide social science explanations for the link formation.

There are few network formation papers that have attempted to account for heterogeneity of agent without losing micro-level interpretability. A study on ecological networks by McKane and Drossel utilized a similar approach, wherein agents are represented by a small number of attributes among a large attribute pool²⁸. However, this work does not directly estimate the latent variables for networks of agents. More abstractly, our method is also reminiscent of mixed membership stochastic blockmodels where agents respectively follow a probability distribution of membership within several communities²⁹. However, probabilistic membership models typically do not seek to uncover economic and sociological mechanisms and the dynamics of network formation. We extend these previous works to the estimation of agent characteristics and network link formation using observed network data. In addition, we want to incorporate a more complex but interpretable inter-agent exchange utility function, by modeling both exchange benefits and coordination costs arising from the differences among agents.

Furthermore, an important question rarely studied in literature is the trade-off between coordination costs and exchange benefits. On the one hand, the coordination between two dissimilar agents incurs higher coordination costs than between two similar agents³⁰, a relationship which encourages homophily, i.e., the tendency to interact more with agents who have shared characteristics³¹. On the other hand, the rationale of exchange benefits

comes from welfare economics: agents have different endowments and their preferences drive different agents to interact and exchange endowments³². The exchange nature therefore encourages heterophily, i.e., the tendency to interact with dissimilar individuals³³. Empirical studies have found that heterophily exists in various scenarios^{34,35}, and that complimentary heterophily between two agents sometimes bring more mutual benefits than homophily³⁶. However, most prior studies of social network formation consider either only coordination costs and homophily^{22,37,38} or only social exchange benefits and heterophily^{39–41}, rather than an integration of exchange and coordination as we do in this paper. The trade-off between exchange benefits and coordination costs is also reminiscent of the identity-diversity balance in the organizational performance literature^{42,43}.

In this paper, inspired by the network embedding techniques, we develop a social network formation model using representation learning methods for heterogeneous agents; to retain the interpretability, we maintain the inter-agent micro-structure characteristics of most agent-based models and the macro-level structures that are the focus of sociology. In our model, agents are characterized by vectors, called their endowment vectors; agents maximize their utility by having link formation driven by comparing their own endowment vectors with those of others. Importantly, we take an economic view of human networks, which considers link formation to be driven by the trade-off between the benefit of exchanges⁴⁴ among individuals with different endowments against the coordination costs due to differences in some other dimensions of endowments. We apply optimization methods to ascertain the endowment vectors of all agents from empirical social networks. The effectiveness of this method is validated by prediction tasks of link formation and individual characteristics. Subsequently, the agent-based models derived from empirical data are evaluated in terms of their micro- and macro-level behavior, compared with the behavior of human networks. Abstractly, we model link formation as a reaction-diffusion system, a framework found in many biological systems.

Results

A game theoretical model. Endowment is a well-known and useful concept in microeconomic theory³², for example, fundamental theorems of welfare economics are based on agent exchanging endowments. In our model, an endowment vector could potentially represent all of the features (assets, abilities, capacities, qualities, etc.) that each agent possesses, and are treated as fixed, invariant characteristics of the agent. We do not consider the situation where endowments are dynamic in this study. Since we limit the dimensionality of endowment vectors, similar to network embedding algorithms (see Methods), each dimension does not necessarily have a specific meaning, but may be a combination of many attributes of an individual.

Agents establish social ties according to the comparison between their endowments. If we assume that there are K dimensions of endowments in a society, each agent has a K -dimensional endowment vector w . Note that dimensions may be mutually correlated; for example, in the Karate club network, leaders and followers have high values in their respective dimensions, and these two dimensions should be negatively correlated. We constrain the first and second moments of each dimension ($\mathbf{W}_{i,k}$) to be zero and one, respectively, for computational simplicity.

We assume the utility function of agent i is only determined by agent i 's neighbors' endowment vectors. We define the utility function $U_i : 2^{\mathcal{I}\{i\}} \rightarrow \mathbb{R}$ for all i , as Eq. (1). The argument S is the potential neighbors, denoting an arbitrary subset of all agents except i herself, i.e., $\mathcal{I}\{i\}$. Each agent i selects her neighbor set S

by maximizing her utility function U_i . U_i is composed of two terms, the benefits of exchange (F_i) and the costs of coordination (G_i):

$$U_i(S; \mathbf{W}, \mathbf{b}, \mathbf{c}) = \underbrace{F_i(S; \mathbf{W}, \mathbf{b})}_{\text{benefits of exchange}} - \underbrace{G_i(S; \mathbf{W}, \mathbf{c})}_{\text{costs of coordination}}, \quad \forall S \subset \mathcal{I}\{i\}. \quad (1)$$

Let S_i^* be the optimal neighbor set for i . We define the marginal utility that j brings to i as:

$$\Delta u_i(j) = \begin{cases} U_i(S_i^*; \mathbf{W}, \mathbf{b}) - U_i(S_i^* / \{j\}; \mathbf{W}, \mathbf{b}), & \text{if } j \in S_i^*; \\ U_i(S_i^* \cup \{j\}; \mathbf{W}, \mathbf{b}) - U_i(S_i^*; \mathbf{W}, \mathbf{b}), & \text{if } j \notin S_i^*. \end{cases} \quad (2)$$

In this study, we are focused on specific forms for F_i and G_i and, consequently, for U_i . For the costs of coordination, agent i 's cost incurred by agent j is measured by the difference between \mathbf{w}_j and \mathbf{w}_i .

$$G_i(S; \mathbf{W}, \mathbf{c}) = \sum_{i \in S} g(\mathbf{w}_j, \mathbf{w}_i, \mathbf{c}) = \sum_{i \in S} \|\mathbf{c} \circ (\mathbf{w}_j - \mathbf{w}_i)\|_2. \quad (3)$$

“ \circ ” denotes element-wise multiplication. $\|x\|_2$ denotes ℓ_2 norm. Note that the costs are symmetric, i.e., $\|\mathbf{c} \circ (\mathbf{w}_i - \mathbf{w}_j)\|_2 = \|\mathbf{c} \circ (\mathbf{w}_j - \mathbf{w}_i)\|_2$. The costly scaling parameter, c_k , measures the importance of k -th dimensions on the costs. A higher c_k will amplify the difference between i and j 's endowment vectors on the k -th dimension ($w_{jk} - w_{ik}$). This term encourages homophily: dissimilar pairs have to suffer from high coordination costs before forming a link.

For F_i , we propose the following form:

$$F_i(S_i^*; \mathbf{W}, \mathbf{b}) = \sum_{j \in S_i^*} \sum_{k=1}^K b_k \max(w_{jk} - w_{ik}, 0). \quad (4)$$

Intuitively, $w_{jk} - w_{ik}$ measures the “advantage” of agent j on the k -th dimension over agent i . As we do not want negative benefits, we consider the benefit on the k -th dimension is zero if $w_{jk} - w_{ik} < 0$. In deep learning, $\max(x, 0)$ is called the “ReLU” function. TensorFlow⁴⁵, a machine learning programming library, provides methods to optimize functions that contain ReLU functions. Similar to c_k , the beneficial scaling parameter b_k measures how beneficial the k -th dimension is. This term indicates that when an agent is high in several dimensions, she could bring high benefits to others. Therefore, other agents are inclined to link to her. However, she does not necessarily reciprocate every link because, for example, when she is higher in every dimension than others, she will not benefit from others in any dimension. Note that for simplicity, we do not consider comparative advantages in this paper. In addition, this term encourages heterophily: agents whose expertises are complimentary have high potential benefits for link formation. Therefore, in this specific form, we have

$$\Delta u_i(j) = \sum_{k=1}^K b_k \max(w_{jk} - w_{ik}, 0) - \|\mathbf{c} \circ (\mathbf{w}_j - \mathbf{w}_i)\|_2 \quad (5)$$

There are of course many other variations for the functional form (Eq. (1)). For example, we can let F_i non-separable in terms of the neighbor set S , e.g., $F_i(S) = \frac{1}{|S|} \sum_{j \in S_i^*} \sum_{k=1}^K b_k \max(w_{jk} - w_{ik}, 0)$.

The intuition is that when one agent has many neighbors, the benefit brought by each neighbor decreases; Do et al. provide a good example of a decreasing marginal utility⁴⁶. However, this functional form indicates that $\Delta u_i(j)$ depends on the neighbor set S , which leads to a time-consuming combinatorial optimization in the learning process; specifically, when the learning algorithm chooses S_i^* , it may need $\mathcal{O}(N2^N)$ computations for the utility functions, which is computationally infeasible for even a small-scale network. This is thus beyond the scope of this paper. We can also change G_i into other norms, such as ℓ_1 norm, or change F_i into a smoother version of $\max(x, 0)$, but these changes do not significantly affect the results in the later sections, as shown in Supplementary Note 9. Therefore, we concentrate on this specific form in later sections (Eq. (5)).

In network game theory, pairwise stability²⁰ refers to the situation where no increased marginal utility can be brought to both agents of an unconnected pair, and no increased marginal utility can be brought to any agents who want to drop their neighbors. Following the definition, we derive the conditions when pairwise stability in undirected networks is satisfied. The proof is straightforward and can be found in Supplementary Note 1.

Proposition 1. An undirected network ($\mathcal{G} = (\mathcal{V}, \mathcal{E})$) implied by neighbor sets S_i^* , $i = 1, 2, \dots, N$ is pairwise stable, if the following conditions are satisfied:

1. if $j \in S_i^*$, then $i \in S_j^*$;
2. $\forall j \in S_i^*$, $\Delta u_i(j) \geq 0$;
3. $\forall j \notin S_i^*$, $\min(\Delta u_i(j), \Delta u_j(i)) < 0$.

Learning endowments. We have established a model for social network formation with many parameters and latent variables. Before we examine the proprieties of the model, we have to assign values for the unknown variables, including the endowment vectors (\mathbf{W}), and scaling parameters (\mathbf{b} and \mathbf{c}). To equip our model with the capability of fitting real-world networks, we learn the endowment vectors using the observations of real-world networks, by assuming real-world networks are at or close to pairwise stability.

Let $\mathcal{L}(\mathbf{b}, \mathbf{c}, \mathbf{W} | D)$ be the loss function that we want to minimize. The definition of $\mathcal{L}(\mathbf{b}, \mathbf{c}, \mathbf{W} | D)$ is reported in Supplementary Note 3. Then we solve the optimization problem in Eq. (6).

$$\begin{aligned} \text{Minimize}_{\mathbf{b}, \mathbf{c}, \mathbf{W}} : & \quad \mathcal{L}(\mathbf{b}, \mathbf{c}, \mathbf{W} | D) \\ \text{Subject to :} & \quad b_k \geq 0, \forall k = 1, 2, \dots, K \\ & \quad c_k \geq 0, \forall k = 1, 2, \dots, K \\ & \quad \sum_{i=1}^N w_{ik} = 0, \forall k = 1, 2, \dots, K \\ & \quad \|\mathbf{W}_{:k}\|_2^2 = N, \forall k = 1, 2, \dots, K \end{aligned} \quad (6)$$

The constraints that b_k and c_k should not be less than 0 are required by the properties of our model. The constraint for the mean of each dimension is to limit the number of equivalent solutions, so that the optimizer could typically find a better solution. The constraint of $\mathbf{W}_{:k}$ is to guarantee that the standard deviation of each dimension is approximately 1, so that the values of \mathbf{b} and \mathbf{c} are comparable across dimensions.

As $\mathcal{L}(\mathbf{b}, \mathbf{c}, \mathbf{W} | D)$ is nonlinear and non-convex (dimensions are interchangeable) with respect to $(\mathbf{b}, \mathbf{c}, \mathbf{W})$, we have to approximate the global optimum by a local optimum. By employing Adam optimizer (an improved stochastic gradient descent method)⁴⁷,

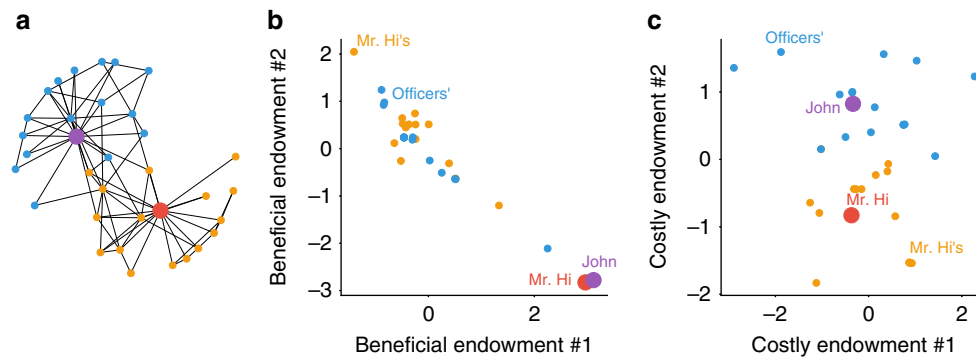


Fig. 1 Illustration of the learned endowments for karate club network. **a** The network structure of karate club network. Mr. Hi, the instructor, is marked red, and other people in his faction are marked orange. John, the student leader or officer, is marked purple, and other people in his faction are marked blue. **b** The first two dimensions of the learned endowment vector for each individual; these two dimensions are only related to exchange benefits so we call them “beneficial endowments”. **c** The last two dimensions of the learned endowment vector for each individual; these two dimensions are only related to coordination costs so we call them “costly endowments”

we are able to learn the local optimum of $\mathcal{L}(\mathbf{b}, \mathbf{c}, \mathbf{W}|D)$; Adam optimizer is good at deriving satisfying local optima when solving nonlinear and non-convex problems. To obtain a solution that approximates the global optimum, we start from many randomly selected initial points and then analyze the results of the multiple runs to find the parameters that generate the smallest loss and therefore the best link fitting performance. Technical details, including the definition of \mathcal{L} and methods that assist learning, are presented in Supplementary Note 3.

Validation of learning. Here we show that we have learned meaningful endowment vectors from empirical networks. In particular, we first use a toy example—Zachary’s karate club network⁴⁸ to illustrate the learned results. We then validate the effectiveness of our model and learning method by showing their performance at fitting link formation and predicting individual characteristics for a variety of large-scale social networks: a synthetic network where two types of agents exchange, a Trade network among countries, a movie collaboration network, a Company communication network, and the Andorra network, which is a nationwide mobile phone network (see Methods).

We start with a toy example to illustrate both the rationale of the present model and the effectiveness of learning performance. Because of a conflict between an instructor (Mr. Hi) and a student officer (John), the social network of Zachary’s karate club is polarized into two factions (Fig. 1a). We set $K=4$ and the first two dimensions as “beneficial endowments” and the last two dimensions “costly endowments” (Methods section) because it is more convenient for visualization if the numbers of beneficial and costly dimensions are both even. Note that $K=4$ is not necessarily the optimal dimensionality and here we did not add a regularization term (Supplementary Note 3) for this result; however, we also show in Supplementary Note 8 that $K=4$ is a reasonable (almost optimal) selection.

Panels b and c in Fig. 1 plot the values of the learned endowments of individuals in Zachary’s karate club. In panel b, both Mr. Hi and John are high in dimension #1 and low in dimension #2, while the rest are generally low in endowment dimension #1 and high in dimension #2. We interpret this result as the tendency of exchanges between instructors and students: dimension #1 represents the professional skill of karate and leadership in their factions; endowment #2 represents the willingness to learn Karate. As for costly endowments (panel c), we find that dimension #4 corresponds to the faction to which each individual belongs: Mr. Hi and his followers (orange) have

Table 1 Learning results and link fitting performance of learned endowment vectors

Dataset	$ \mathcal{V} $	$ \mathcal{E} $	K^*	K_{bnf}^*	K_{cst}^*	AUC
Karate	34	78	3	2	1	98.48%
Trade	100	703	5	3	2	96.85%
Synthetic	2500	14,453	4	2	2	99.92%
Company	1984	12,751	11	4	7	98.70%
Movie	2788	10,399	7	4	3	96.08%
Andorra	32,829	513,931	15	8	7	94.76%

$|\mathcal{V}|$ denotes the number of nodes and $|\mathcal{E}|$ denotes the number of edges. K^* , K_{bnf}^* , and K_{cst}^* represent the optimal number of dimensions, beneficial dimensions, and costly dimensions, respectively

values generally higher than 0 while John and his followers (blue) are generally lower than 0. Dimension #4 can be explained as the individual’s identification with the two factions. We interpret cost endowment #3 as other unobserved characteristics that might influence the interactions between individuals, such as the time and frequency to participate in club activities. We also illustrate the learning results for the Trade and Synthetic datasets graphically in Supplementary Note 4.

Because our goal is to use the learned endowment vectors to further analyze the micro- and macro- patterns of the network, we learn the endowment vectors by using all the information (the links) of the network. Therefore, rather than split the input links into training and test sets, we use all the links as the input. A potential concern is that we might “overfit” the network by using a large K ; we partially address this concern by introducing the regularization term \mathcal{L}_{reg} as mentioned in Supplementary Note 3. We use $\Delta u_i(j)$ as the predictor and AUC (area under the curve) as the measurement for the fitting performance. AUC trades off between true positive and false positive rates, and serves as a fair measure when there is a strong imbalance between positive and negative samples. By using an approach provided in Supplementary Note 3, we obtain the optimal dimensionality (K) and the optimal number of beneficial and costly endowments (K_{bnf} and K_{cst} ; see their definitions in Methods).

As shown in Table 1, our model is able to obtain very good fits to the input networks. For all datasets, the AUC of link fitting is over 94%. Moreover, we demonstrate that for all datasets, it is necessary to incorporate both the benefit and the cost terms into

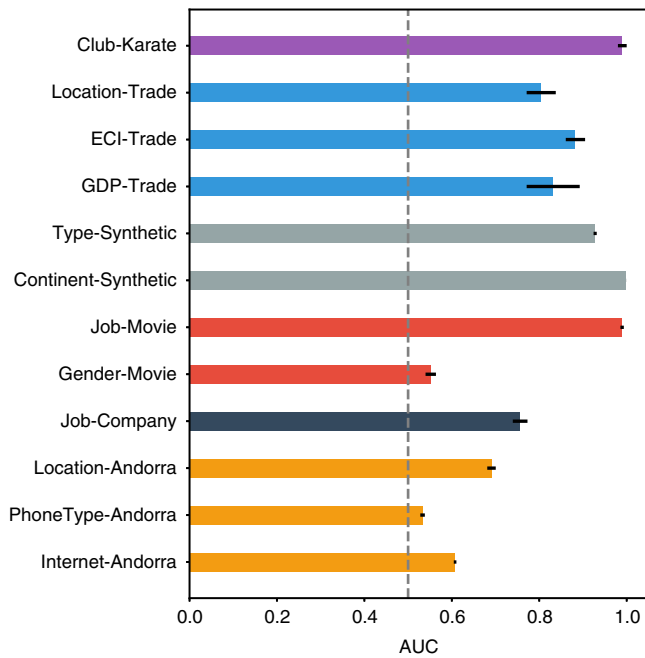


Fig. 2 Prediction performance of the learned endowments. We use support vector machine as the classifier. The baseline, random guess algorithm, is indicated by the dashed line. Error bars represent the standard errors for the average AUCs in five random splitting of training-test sets

the utility functions (i.e., $K_{bnf} > 0$ and $K_{cst} > 0$). This finding highlights the importance of integrating both exchange effects and coordination costs into the link formation mechanisms. Other technical details, including learning curves and the performance on all the dimensions, are presented in Supplementary Note 3.

Although our goal is not to design a network embedding algorithm that outperforms the state-of-the-art algorithms, it is interesting to examine our model’s ability to predict individual characteristics as a network embedding algorithm. If the learned endowments have a decent predictive power for individual characteristics, we can then believe that we have effectively learned the endowment vectors, which can be used for further analysis such as agent-based modeling. We extract characteristics that are not directly relevant to nodes’ ego network attributes (see Supplementary Note 2 for a full list). We split the nodes and their learned endowment vectors into training (75%) and test (25%) sets. We use support vector machine (SVM) and k -nearest neighbors algorithm (k -NN) to train the classifiers, and use cross-validation to tune the classifiers’ hyperparameters.

As shown in Fig. 2, the learned endowment vectors can well predict most individual characteristics by SVM. Note that k -NN has similar results in Supplementary Note 5. This result shows that our model can encapsulate the latent features of agents. It is important to highlight that individual characteristics might not be fully reflected in the network; therefore, neither network embedding algorithms nor the present model can guarantee high AUCs for all prediction tasks. However, the learned endowment vectors in fact contain more information than the presented agent features; therefore, they could predict agent characteristics that are not used in this work, e.g., preferences of movie genres.

The accuracy at estimating agent characteristics beyond the input data could be because they are important either in coordination costs (e.g., locations) or exchange benefits (e.g., collaboration between cast members and directors). Some characteristics may

have both exchange effects and coordination costs: for example, in a company, subordinates mostly communicate with each other (low coordination costs), but would also interact with their managers occasionally (exchange benefits).

We also compare our results with a network embedding algorithm, DeepWalk²⁴, with the same number of dimensions and therefore the same degree of freedom (Supplementary Note 5). Recall that network embedding methods are designed only for dimension reduction; they therefore do not provide economic or sociological insights about the network. Algorithmically, DeepWalk uses an energy function that considers only similarity and not the benefit that can flow from exchanges between agents with very different endowments. Consequently, as might be expected, when our model is compared to DeepWalk, we have better performance if the predicted characteristics are explicitly implied by exchange effects. However, for characteristics explicitly implied by low coordination costs between similar people, the performance of the present model is somewhat lower than that of DeepWalk, probably because DeepWalk considers the similarity between neighbors spanning multiple hops. In sum, the ability to predict agent characteristics shows that our model has learned useful information implicit in the network, and that this implicit information can be used for further agent-based modeling.

Agent-based modeling. We next analyze the properties of the model as an agent-based model. Because of the high degree of freedom of the present model, any manually input distributions of \mathbf{W} , \mathbf{b} , and \mathbf{c} may appear too arbitrary and do not reflect any real-world situation. We therefore use the learned endowments and parameters as the input to study both micro- and macro-level properties of this model. Our model exhibits many complex and well-known social phenomena, suggesting that these phenomena could be caused by the simple mechanisms of exchange benefits and coordination costs among heterogeneous agents.

At the micro level, an interesting question is how an agent’s endowments will affect their ego networks. In particular, we consider two variables for agents based on our model. The first variable is a quantitative measure of social status that we call “social power”

$$\text{social power}(i) = \mathbf{b} \cdot \mathbf{w}_i. \tag{7}$$

Social power means “the potential for social influence”⁴⁹, or the potential benefits that one could bring to the other. Recall that b_k measures how beneficial the k -th dimension is. w_{ik} is the i -th agent’s value on the k -th dimension. As $b_k \times w_{ik}$ increases, i is more likely to benefit others on the k -th dimension. Therefore, it is sensible to represent an agent’s social status by the dot product of \mathbf{b} and \mathbf{w}_i . Therefore, the definition of this variable is consistent with the concept, social power. The utility of this social power for social exchange leads naturally to the formation of a network structure, which is often described as hierarchical, especially within the surrounding homophilic group.

The second variable is “social exclusion”, which measures the extent to which an agent is marginalized⁵⁰:

$$\text{social exclusion} = \|\mathbf{c} \circ \mathbf{w}_i\|_2. \tag{8}$$

Recall that we have constrained the means for all dimensions to be 0. If an agent has a large absolute value on some dimension, she is believed to be on the margin of that dimension because a higher cost is needed when she links to another arbitrary person.

We are interested in the correlation between the social power or social exclusion and statistics of their ego networks (i.e., degree

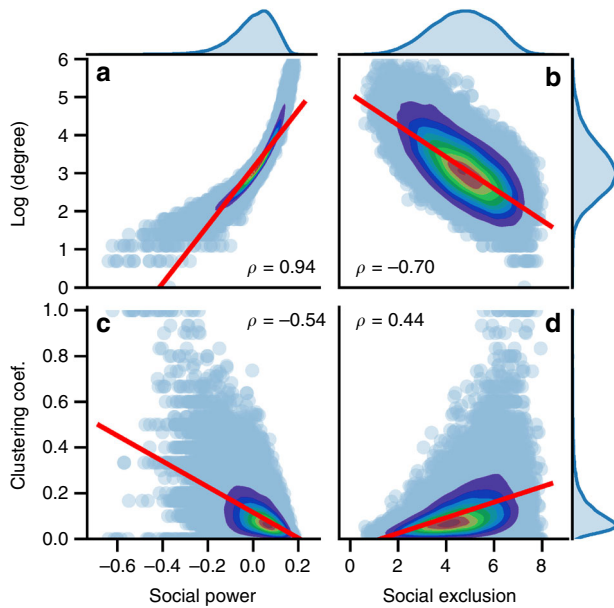


Fig. 3 Correlations between endowment-related variables and ego networks statistics for the Andorra dataset. **a** Social power versus degree. **b** Social exclusion versus degree. **c** Social power versus clustering coefficient. **d** Social exclusion versus clustering coefficient. Colors represent the density of data points, where red indicates dense areas and purple indicates sparse areas. All data points are plotted as light blue dots. ρ denotes Pearson correlation coefficients. All Pearson correlation coefficients are significant at level $p < 0.001$

and clustering coefficient). The results of the Andorra dataset is presented in Fig. 3, and similar results for other datasets are reported in Supplementary Note 6. We find that “social power” is strongly positively correlated with degree, while “social exclusion” is strongly negatively correlated with degree. This finding is consistent with the implication of the proposed model: people with high (beneficial) endowments can potentially benefit others to a greater degree; people on the margin of the society have fewer opportunities to interact with others. More interestingly, we examine the correlations between social power or exclusion and the clustering coefficients for the nodes. A high clustering coefficient means that the agent’s neighbors are closely connected, and therefore indicates that the agent’s neighbors might lack diversity. We find that people have lower clustering coefficients on the network if they have higher social power or lower social exclusion; that is, high status (power) people have more diverse social networks, a well-known and important aspect of human networks.

The proposed model can also predict macro-level dynamics of networks. As an illustration, we are focused on the impact of the systematic change of cost scaling parameters \mathbf{c} (i.e., reducing \mathbf{c} to $\mathbf{c}' = (1 - \alpha)\mathbf{c}$, $\alpha \in [0, 1]$) on the macro statistics of the social network. Decreases in coordination costs are typically caused by advances in information technology (e.g., the Internet) or transportation (e.g., a new railway). We then employ agent-based modeling according to the learned endowment vectors and utility functions to reconstruct the empirical social networks (see Supplementary Note 7 for the approach). Finally, we compute density, average clustering coefficient, average shortest path in the giant component, and interaction diversity (defined as Eq. (9)), where \mathcal{E} represents the edge set of the network, and \mathbf{c} is the value after being reduced. Note that here we do not change the relative ratios among c_k ($1 \leq k \leq K$); it is therefore sensible to incorporate

the \mathbf{c} into Eq. (9) after being normalized by $\|\mathbf{c}\|_2$.

$$\text{interaction diversity} = \frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} \frac{\|\mathbf{c} \circ (\mathbf{w}_i - \mathbf{w}_j)\|_2}{\|\mathbf{c}\|_2} \quad (9)$$

Figure 4 shows the impact of reducing \mathbf{c} on the macro statistics of all networks. We find that as the cost scaling parameters \mathbf{c} decrease, the density significantly increases while clustering coefficient does not increase much. This indicates that the decrease in coordination costs (e.g., adoption of the Internet) results in more links, and increases social cohesion or balance⁵¹, i.e., the connectivity between one’s neighbors. The decreasing trend of shortest paths between pairs reveals that the decrease of the coordination cost could diminish the power of social hierarchy. The trend of interaction diversity indicates that the decrease of coordination costs leads to greater connections between more dissimilar individuals. These synthetic findings indicate that the coordination costs’ reduction, usually caused by technology advances, results in a society with less hierarchy and more opportunities for social connection, especially for dissimilar people.

Discussion

Inspired by network embedding methods that represent agents by vectors, this study also applies vector representations for heterogeneous agents, referred to as their “endowment vectors”. Our model is more interpretable than network embedding algorithms because we can economically and sociologically explain the link formation mechanism, by the trade-off between the exchange benefits and coordination costs among agents. We learned the endowment vectors from empirical network data, which can be used to predict a variety of other agent properties, and to demonstrate inter-agent network characteristics such as social status and diversity that are well-known from social science literature.

In particular, we highlight the necessity of trading off between beneficial exchange effects and coordination costs. Most link formation models use only one or the other. We show that we can effectively learn the representations for agents from empirical networks by optimization methods that incorporate these trade-offs, without explicitly modeling social status, hierarchy, or the dynamics of social networks. This result suggests that many characteristics that are described in the social science literature are due to the trade-off between coordination costs and exchange benefits, rather than being fundamental effects or biases.

There are several interesting future directions based on this work. First, it is intriguing to consider the influence of existing neighbors on the marginal utilities of adding one more neighbor. For instance, the marginal utility of befriending a person should be higher when an ego has 10 friends than when the ego has 100 friends. Incorporating this interaction effect is difficult because this will require combinatorial optimization methods. Second, it is a promising direction to incorporate an indirect effect: the utility of “friends’ friends”. When we befriend a person, we do not only benefit from this person, but also this person’s friends because we obtain useful information from and have small coordination costs with this person’s friends. The indirect effect is reminiscent of several network embedding methods, including DeepWalk, which embed nodes on randomly sampled paths to have similar representations. Finally, we may take into account broader interaction effects such as “reputation”: when people reach out to an ego, the ego may reciprocate a link even if the link does not directly benefit the ego.

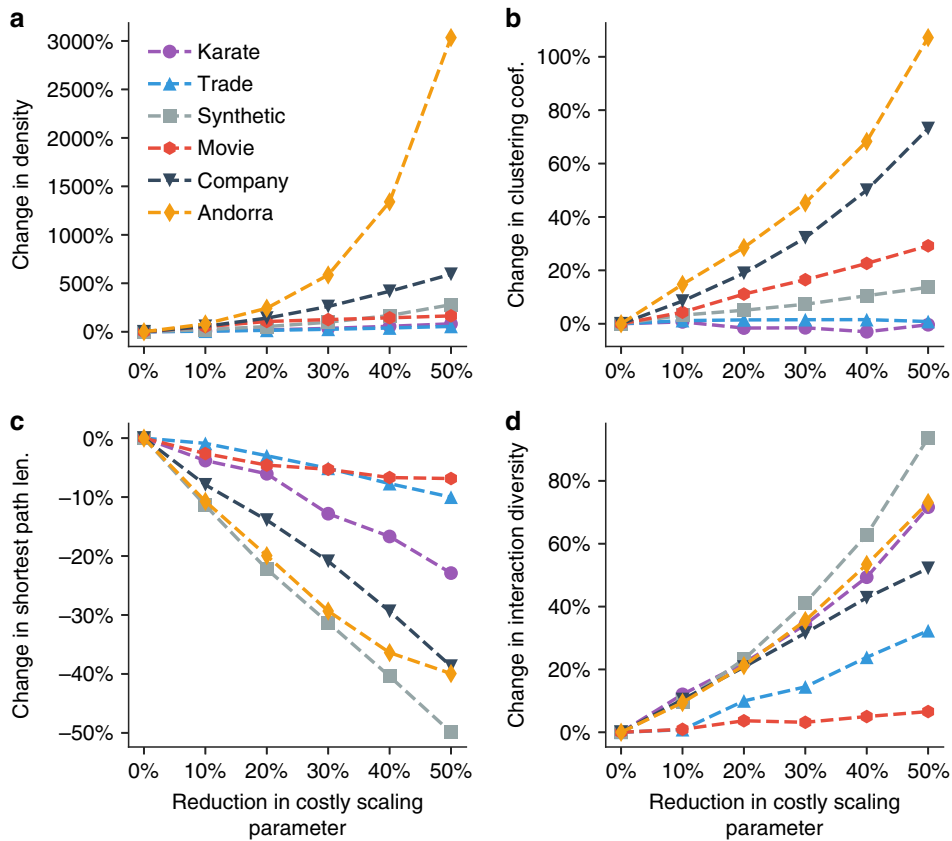


Fig. 4 Impact of reduction in costly scaling parameters on macro-level network statistics. **a** Density. **b** Clustering coefficient. **c** Average of shortest path lengths. **d** Interaction diversity. X-axis represents the reduction of costly scaling parameters; y-axis represents the dynamics of the macro-level network statistics when costly scaling parameters are reduced

Methods

Problem setup. Let $\mathcal{I} = \{1, 2, \dots, N\}$ be a group of N and potentially connected agents indexed by i (or j). Let K be the dimensionality of endowments that drives the formation of the social network of the group, indexed by k . Each agent has a latent endowment vector $\mathbf{w}_i = (w_{i1}, \dots, w_{iK})^T$, with each dimension indicating an aspect of the individual's attributes. Let $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_N)^T$. We observe all edges among the N agents. Let D be a set of $N \times N$ adjacency matrices among agents in all periods. D_{ij} is binary ($\{0,1\}$). $D_{ij} = 1$ if there is an edge from i to j , and $D_{ij} = 0$ otherwise. For the convenience of showing pairwise stability, the study is restricted to undirected graphs, i.e., $D_{ij} = D_{ji}$.

Agents make rational choices by comparing their endowment vectors with potential friends. Agents maximize their utility functions ($U_i : 2^{\mathcal{I}/\{i\}} \rightarrow \mathbb{R}$ for each i) dependent on the differences between their endowment vectors and all possible candidates (all other agents). U_i is also parameterized by \mathbf{W} , \mathbf{b} , and \mathbf{c} . $\Delta u_i(j)$ is the marginal utility that j brings to i . We therefore predict D_{ij} by $\Delta u_i(j)$.

Data description.

- **Andorra.** We collected the nationwide call detail records in Andorra from July 2015 to June 2016. Utilizing the country code, we filtered out all non-citizens, leaving 32,829 citizens with at least one call interactions with another. If the (i, j) had at least one effective call (duration greater than 0 s), we set $D_{ij} = D_{ji} = 1$; otherwise $D_{ij} = D_{ji} = 0$. This process results in 513,931 links. To demonstrate the effectiveness of the learned endowments, we also extracted three characteristics of individuals: phone type, frequent city, and Internet usage. The phone type was identified by the type allocation code, and we classified each type into Apple, Samsung, and others (the distribution of three types is balanced). For each phone number, we employed the last phone type that we observed. Note that type phone is strongly correlated with important individual characteristics such as income. The most frequent city was identified by the cell tower id. We classified each phone number by the location where it shows up most frequently throughout the year, this location is thus likely the work location of the individual (some individuals' work location may be their home). Internet usage was computed by the total duration of cellular data. In the prediction task, we classified Internet usage into high (more than median) and low (less or equal than median). Details of

the datasets, such as statistics of individual characteristics and network degree distribution, are shown in Description in detail in Supplementary Note 2.

- **Movie.** To highlight the exchange effects, we examine a specific type of social network, director-cast movie collaboration network, where a node represents either a movie director or an actor/actress, and an edge between a director i and an actor/actress j represents a collaboration between i and j . $D_{ij} = D_{ji} = 1$ means that i and j collaborated at least once; 0 otherwise. Note that the social network is close to a bipartite graph where nodes are partitioned into directors and cast (some people have both cast and director experience). We extracted 3493 movies throughout 2000–2016, and retained individuals with at least five movies within this period, resulting in 160 directors and 2628 cast members, and 10,399 director-cast pairs. To validate the effectiveness of the learned endowments, we extracted two individual characteristics: occupation and gender. For occupation, we labeled an individual as a director if she functioned as a director in more than a half of the movies in which she engaged; cast otherwise. For gender, we collected 1840 males and 761 females and 186 unlabeled.
- **Synthetic.** We manually establish a network of 2500 agents. Agents are indexed by (x, y) ($i = 50x + y$), $0 \leq x \leq 49$, $0 \leq y \leq 49$, $x, y \in \mathbb{N}$. Each agent therefore resides at a unique location on the 50×50 grid, and the agent has a probability of 0.5 to be either type A (e.g., a buyer) or type B (e.g., a seller). Buyers (sellers) are exploring sellers (buyers) in their neighborhood with Manhattan distance ≤ 3 . The network is therefore a bipartite graph where buyers and sellers exchange goods and money. This data generating process results in 14,453 edges. We predict the type and location (divide the plane into four parts) for all agents.
- **Company.** A network of employees in a company where edges represent a call and text communication (MobileD in⁵²). Each employee is labeled as a manager or a subordinate. In total, we have 420 managers and 1564 subordinates, with 12,751 edges among them. In this network, managers are mostly connected with managers and subordinates are mostly connected with subordinates. At the same time, subordinates also interact with their respective managers occasionally. We believe that this dataset should show a trade-off between coordination and exchange; for example, managers and subordinates have exchange effects, and they have lower coordination costs to interact with the same type.

- Trade. We use the 2014 international trade data provided by the United Nations Statistical Division (UN Comtrade Database: <https://comtrade.un.org/>), specifically the cleaned version provided by the BACI team using their own methodology of harmonization⁵³. We created a network of countries, where an edge indicates that the trade value between two countries is >1 billion dollars (for both directions). This process resulted in 100 countries with at least one link, and 703 undirected edges among them. We predict the GDP, economic complexity index (ECI)⁵⁴, and the countries' continents for this dataset.

Details in learning. For computational simplicity and better fitting performance (see Supplementary Note 8), we split the dimensions into “beneficial dimensions” and “costly dimensions”. In Eq. (5), every dimension (say the k -th) can contribute to both benefits and costs if both b_k and c_k are greater than zero. However, it is not difficult to see that if we constrain some dimensions to have zero-valued beneficial scaling parameters ($b_k = 0$) or costly scaling parameters ($c_k = 0$), the dimensionality of the model (K) will increase but the capacity of data fitting will not change. During the learning process, a connected pair (i, j) may result in either an increase in the difference on some beneficial dimension (with $b_k > 0$) or a decrease in the difference on some costly dimension (with $c_k > 0$) between their endowment vectors. Empirically, if both b_k and c_k are positive, these two conflicting effects (to increase or to decrease the utility on the same dimension) would hinder an effective convergence (shown in Supplementary Note 8); we conjecture that this is because we are optimizing a non-linear non-convex loss function. Therefore, we separate the K dimension into K_{bnf} “beneficial dimensions” and K_{cst} “costly dimensions” ($K_{\text{bnf}} + K_{\text{cst}} = K$). By comparing the performances of link fitting for different K_{bnf} and K_{cst} , we select the optimal K_{bnf}^* and K_{cst}^* , and consequently K^* . For simplicity, we let $b_k = 0$, for $k > K_{\text{bnf}}^*$, and $c_k = 0$, for $k \leq K_{\text{bnf}}^*$. $\theta = (b_1, b_2, \dots, b_{K_{\text{bnf}}^*}, c_{K_{\text{bnf}}^*+1}, c_{K_{\text{bnf}}^*+2}, \dots, c_K)$. In Supplementary Note 8, we show empirically that the performances of link fitting and node classifications are worse when we do not split dimensions into beneficial and costly dimensions; and that even when we do not split dimensions, the learning algorithm will lead most dimensions to be either “beneficial” or “costly”, i.e., either b_k or c_k is very close to zero. More details can be found in Supplementary Note 3.

Code Availability. Code is available online: <https://github.com/yuanyu94/endorment>.

Data availability

The network data and individual attributes are available online: <https://github.com/yuanyu94/endorment>.

Received: 4 May 2018 Accepted: 11 October 2018

Published online: 08 November 2018

References

- Newman, M. *Networks: an introduction* (Oxford Univ. Press, Oxford, 2010).
- Wasserman, S. & Faust, K. *Social network analysis: methods and applications*, vol. 8 (Cambridge Univ. Press, Cambridge, 1994).
- Jackson, M. O. A survey of network formation models: stability and efficiency. *Group Formation in Economics: Networks, Clubs, and Coalitions* 664, (11–49. Cambridge University Press, New York, 2005).
- Borgatti, S. P., Mehra, A., Brass, D. J. & Labianca, G. Network analysis in the social sciences. *Science* **323**, 892–895 (2009).
- Rogers, E. M. *Diffusion of innovations* (Simon and Schuster, New York City, 2010).
- Bakshy, E., Rosenn, I., Marlow, C. & Adamic, L. The role of social networks in information diffusion. In *Proc. 21st International Conference on World Wide Web*, 519–528 (ACM, Lyon, France, 2012).
- Fowler, J. H. & Christakis, N. A. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *BMJ* **337**, a2338 (2008).
- Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. The diffusion of microfinance. *Science* **341**, 1236498 (2013).
- Fiorina, M. P. & Abrams, S. J. Political polarization in the american public. *Annu Rev. Polit. Sci.* **11**, 563–588 (2008).
- Aral, S., Muchnik, L. & Sundararajan, A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl Acad. Sci. USA* **106**, 21544–21549 (2009).
- Meyers, L. A., Newman, M. & Pourbohloul, B. Predicting epidemics on directed contact networks. *J. Theor. Biol.* **240**, 400–418 (2006).
- Christakis, N. A. & Fowler, J. H. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357**, 370–379 (2007).
- Bond, R. M. et al. A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295 (2012).
- Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440 (1998).
- Barabási, A.-L. Scale-free networks: a decade and beyond. *Science* **325**, 412–413 (2009).
- Jackson, M. O. & Wolinsky, A. A strategic model of social and economic networks. *J. Econ. Theory* **71**, 44–74 (1996).
- Skyrms, B. & Pemantle, R. A dynamic model of social network formation. *Proc. Natl Acad. Sci. USA* **97**, 9340–9346 (2000).
- Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502 (2006).
- Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
- Jackson, M. O. *Social and economic networks* (Princeton Univ. Press, Princeton, 2010).
- Mele, A. A structural model of dense network formation. *Econometrica* **85**, 825–850 (2017).
- Christakis, N. A., Fowler, J. H., Imbens, G. W. & Kalyanaraman, K. An empirical model for strategic network formation, Preprint at <http://www.nber.org/papers/w16039> (2010).
- Chandrasekhar, A. G. & Jackson, M. O. A network formation model based on subgraphs, Preprint at <https://arxiv.org/abs/1611.07658> (2016).
- Perozzi, B., Al-Rfou, R. & Skiena, S. Deepwalk: Online learning of social representations. In *Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 701–710 (ACM, New York City, USA, 2014).
- Tang, J. et al. Line: large-scale information network embedding. In *Proc. 24th International Conference on World Wide Web*, 1067–1077 (ACM, Florence, Italy, 2015).
- Grover, A. & Leskovec, J. Node2vec: scalable feature learning for networks. In *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 855–864 (ACM, San Francisco, USA, 2016).
- Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. In *Proc. 5th International Conference on Learning Representations* (ACM, Toulon, France, 2017).
- McKane, A. J. & Drossel, B. Models of food web evolution. *Ecological Networks: linking Structure to Dynamics in Food Webs* 223–243 (Oxford Univ. Press, Oxford, 2006).
- Airoldi, E. M., Blei, D. M., Fienberg, S. E. & Xing, E. P. Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9**, 1981–2014 (2008).
- Jackson, M. O. & Xing, Y. Culture-dependent strategies in coordination games. *Proc. Natl Acad. Sci. USA* **111**, 10889–10896 (2014).
- McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001).
- Mas-Colell, A. et al. *Microeconomic theory* (Oxford university press, New York, 1995).
- Rogers, E. M. & Bhowmik, D. K. Homophily-heterophily: relational concepts for communication research. *Public Opin. Q.* **34**, 523–538 (1970).
- Johnson, N. F. et al. Human group formation in online guilds and offline gangs driven by a common team dynamic. *Phys. Rev. E* **79**, 066117 (2009).
- Kimura, D. & Hayakawa, Y. Coevolutionary networks with homophily and heterophily. *Phys. Rev. E* **78**, 016103 (2008).
- Alpert, M. I. & Anderson, W. T. Optimal heterophily and communication effectiveness: some empirical findings. *J. Commun.* **23**, 328–343 (1973).
- Boguná, M., Pastor-Satorras, R., Daz-Guilera, A. & Arenas, A. Models of social networks based on social distance attachment. *Phys. Rev. E* **70**, 056122 (2004).
- Currarini, S., Jackson, M. O. & Pin, P. An economic model of friendship: homophily, minorities, and segregation. *Econometrica* **77**, 1003–1045 (2009).
- Cook, K. S. & Yamagishi, T. Power in exchange networks: a power-dependence formulation. *Soc. Network* **14**, 245–265 (1992).
- Friedkin, N. E. An expected value model of social power: predictions for selected exchange networks. *Soc. Network* **14**, 213–229 (1992).
- Kleinberg, J. & Tardos, É. Balanced outcomes in social exchange networks. In *Proc. 40th Annual ACM Symposium on Theory of Computing*, 295–304 (ACM, Victoria, Canada, 2008).
- Watson, W. E., Kumar, K. & Michaelsen, L. K. Cultural diversity’s impact on interaction process and performance: comparing homogeneous and diverse task groups. *Acad. Manag. J.* **36**, 590–602 (1993).
- Hong, L. & Page, S. E. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc. Natl Acad. Sci. USA* **101**, 16385–16389 (2004).
- Page, S. E. *The difference: how the power of diversity creates better groups, firms, schools, and societies* (Princeton Univ. Press, Princeton, 2008).
- Abadi, M. & TensorFlow, A. A. B. P. Large-scale machine learning on heterogeneous distributed systems. In *Proc. 12th USENIX Symposium on Operating Systems Design and Implementation*, 265–283 (USENIX, Savannah, USA, 2016).

46. Do, A.-L., Rudolf, L. & Gross, T. Patterns of cooperation: fairness and coordination in networks of interacting agents. *New J. Phys.* **12**, 063023 (2010).
47. Kingma, D. & Ba, J. Adam: a method for stochastic optimization. In *Proc. 3rd International Conference on Learning Representations* (ACM, San Diego, USA, 2015).
48. Zachary, W. W. An information flow model for conflict and fission in small groups. *J. Anthropol. Res.* **33**, 452–473 (1977).
49. French, J. R., Raven, B. & Cartwright, D. The bases of social power. *Class. Organ. Theory* **7**, 311–320 (1959).
50. Strauss, R. S. & Pollack, H. A. Social marginalization of overweight children. *Arch. Pediatr. Adolesc. Med.* **157**, 746–752 (2003).
51. Cartwright, D. & Harary, F. Structural balance: a generalization of heider's theory. *Psychol. Rev.* **63**, 277 (1956).
52. Tang, J., Lou, T., Kleinberg, J. & Wu, S. Transfer link prediction across heterogeneous social networks. *ACM Trans Inf Syst* **9**, Article 43 (2010).
53. Gaulier, G. & Zignago, S. Baci: international trade database at the product-level (the 1994–2007 version) <https://ideas.repec.org/p/cii/cepiddt/2010-23.html> (2010).
54. Hidalgo, C. A. & Hausmann, R. The building blocks of economic complexity. *Proc. Natl Acad. Sci. USA* **106**, 10570–10575 (2009).

Acknowledgements

This work was supported in part by the MIT Trust Data Consortium and the King Abdulaziz City for Science and Technology. We thank Kent Larson, ActuaTech, and the City Science Group at MIT Media Lab for their support of the Andorra dataset. We thank Abdullah Almaatouq, Xiaowen Dong, Eaman Jahani and Yan Leng for their comments.

Author contributions

Y.Y., A.A., and A.S.P. conceived the present idea. Y.Y. and A.A. planned the experiments. Y.Y. designed the model and performed the experiments. Y.Y., A.A., and A.S.P. analyzed the results. Y.Y. wrote the paper with input from A.A. and A.S.P.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-07089-x>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018