Building Relationships in Evolving Networks: Exploitative Versus Exploratory Strategies

Yang Chen[1], Bo Yan[1], and Jiamou Liu[2]

[1] Beijing Institute of Technology, Beijing, China
[2] University of Auckland, Auckland, New Zealand
jiamou.liu@auckland.ac.nz

Abstract. Interpersonal relations facilitate information flow and give rise to positional advantage of individuals. We ask the question: *How would an individual build relations with members of a dynamic social network in order to arrive at a central position in the network?* More formally, we propose the dynamic network building problem. Two strategies stand out to solve this problem: The first directs the individual to exploit its social proximity by linking to nodes that are close-by. The second links the individual with distant regions of the network. We test and compare these two strategies with edge- and distance-based cost metrics. Experiments over standard dynamic network models and real-world data sets reveal that the exploitative strategy, despite relying on restricted parts of the network, often gives comparable or superior results. We then present and test ways that combine these two strategies.

Keywords: dynamic social networks, interpersonal ties, network evolution, centrality

1 Introduction

The network of social relations entails important properties of individuals. Take, as an example, the structural construct of centrality [27]. Much has been revealed about the correlation between centrality and social statues [8, 9, 22, 23]. By occupying a more central position in the social network, an individual may exercise more control over the flow of information, accessing diverse knowledge and skills, and hence gaining a higher positional advantage [32]. Exploiting this principle, individuals may cultivate relationships with others towards improving their social statues [10]. One famous example is the House of Medici, who rose to prominence in 15th century Florence through intermarriage with other noble families [29]. Another example is Moscows growing statues in 12-13th century Russia thanks to trade relationships with other towns [30].

Imagine that an individual tries to embed herself at the center of a social network through forming new ties. From a structural perspective, this individual needs to choose a set of members to build links with[1]. To this end, the individual

[1] Here we put aside issues such as attitude, personality, and individual preferences, and focus on a structural perspective of network building.

may adopt an *exploitative* or an *exploratory* strategy: The former ensures that the individual exploits existing interpersonal ties and links to those that share a common social proximity; On the contrary, the latter allows the individual to explore far and bridge diverse parts of the network. A natural question arises as to which strategy is more suitable. Moreover, social networks in real life are rarely static, but rather, they constantly evolve with time. Thus the question has an extra layer of complexity: *How to incrementally build relationships in a network to gain positional advantage while the network is evolving?*

To attempt this question, we should settle several issues: Firstly, we need a notion that reasonably reflects positional advantage; here centrality metrics may be of use. Secondly, relation building costs time and effort; one needs to quantify such costs. Thirdly, one needs models on how a social network evolves.

Contribution (1) In this paper, we propose the problem of *dynamic network building (DNB)*. The input to the problem consists of a connected graph G that undergoes a sequence of updates. The problem asks for a plan that builds edges incrementally between a node v and other nodes so that v gains centrality as G evolves. (Sec. 2) **(2)** To solve this problem, we define exploitative and exploratory strategies and present heuristics to realize each strategy. (Sec. 3) **(3)** We compare the heuristics over various evolution models of social networks and real-world networks. Exploration often builds less number of new links, while the exploitative strategy produces better results when other factors, such as distance and embeddedness is considered (Sec. 4). **(4)** Lastly, we propose and evaluate ways that combine the exploitative and exploratory strategies (Sec. 5).

This work is meaningful in the following ways: Firstly, the exploitative strategy resembles the acquaintance process in real-life: It is usually easier to acquaint with those who are close to our own social circles than with those who are far away [26]. Our results demonstrate that this intuition may give us effective means to build relationship in a network. Secondly, while exploratory strategy tends to improves centrality more quickly, exploitative strategy results in a much higher embeddedness, which leads to stronger ties and a platform of trust [11]. Thirdly, the tradeoff between exploration and exploitation has been a recurring theme in artificial intelligence and knowledge management [28, 5, 33]. This work discovers an incarnation of this tradeoff in the context of social networks and ways to mix the two strategies to produce suitable strategies for relationship building.

Related works. The establishment of interpersonal ties has been a major problem in social network analysis. Granovetter's pioneering work contrasts ties having high embeddedness (strong ties) with ties that bridge two otherwise disjoint social circles (weak ties); while embeddedness reflects important dimensions such as trust, commitment and solidarity [17], bridges are important to the exchange of knowledge and ideas [15]. We extends this discussion to study strategies for building different types of ties. Network building (NB) has been studied in [25]. The problem studied in this paper has crucial differences: (a) While NB only operates on static networks, here we focus on evolving networks, which demand the node to be strategic towards future changes. (b) While NB focuses on smallest eccentricity, DNB aims for optimal closeness centrality. (c) DNB considers costs incurred from the distance between the two nodes when forming an edge. A large literature on *strategic network formation* explains tie establishment between rational agents using game theory; these works do not consider stochastic models of network evolution [19]. The exploratory and exploitative strategies of dynamic network building parallel the two modes of network formation in [20]; there, "meeting strangers" means exploratory encounters in the network, and "meeting friends-of-friends" means utilizing existing social circles.

2 Dynamic Network Building Problem

A *social network* is a graph $G = (V, E)$ where V is a set of nodes and E is a set of undirected edges on V of the form uv where $u \neq v \in V$. $\Gamma(u) = \{v \mid uv \in E\}$ denotes the *neighborhood* of u. A *path* (of *length* k) is a sequence of nodes $u_0, u_1, \ldots, u_k$ where $u_i u_{i+1} \in E$ for any $0 \leq i < k$. The (geodesic) *distance* between u and v, denoted by $\mathsf{dist}_G(u, v)$, is the length of a shortest path between u and v. We omit the subscript G writing simply $\mathsf{dist}(u, v)$ when the underlying graph is clear. We also need the following formalism:

- For a node $s \in V$ and $v \neq s$, denote by $G \oplus_s v$ the expanded network $(V \cup \{v\}, E \cup \{sv\})$.
- We assume that the social network G evolves by some (discrete-time) stochastic mechanism, which we define below:

Definition 1. *An* evolution mechanism M *is a function that maps a social network* G *to a probability distribution of social networks* $M(G)$*. Starting at* G*, the network evolves to a sample outcome of* $M(G)$ *in the next time step.*

Imagine v is a node who wants to build relationships in G (let's call v the *newcomer*). We assume that (1) v is a node with few connections in G; (2) v may create edges from itself to nodes in V by paying costs (see below); and (3) v has no knowledge regarding how G may evolve.

Abstractly, one can view the interactions between v and the social network G as a two-player game; the players are v and a player representing the evolution mechanism of the network. At each round, v creates an edge with a node in G (keeping all existing edges).[2] The evolution mechanism then modifies the updated network. Through multiple rounds, v aims to get increasingly integrated into the network. Note that we assume that v functions independently from the evolution mechanism to highlight that v builds edges without prior knowledge of the network evolution mechanism.

Definition 2. *An (ℓ-round)* network building (NB) process *between* v *and* G *consists of a sequence of networks* $G_0 = (V_0, E_0), G_1 = (V_1, E_1), \ldots, G_\ell = (V_\ell, E_\ell)$ *and a sequence of nodes* $s_0 \in V_0, s_1 \in V_1, \ldots, s_{\ell-1} \in V_{\ell-1}$ *such that* $G_0 = G$ *and each network* G_{i+1} *is a sample output of* $M(G_i \oplus_{S_i} v)$.

[2] For simplicity, we assume each round allows v to create at most one new edge. One may easily generalize this setting to allow v to create several edges in a single round.

Definition 3. *An* NB strategy *is a function φ that outputs a node $\varphi(G)$ in a given network G. Any NB process $(G_0, \ldots, G_\ell, s_0, \ldots, s_{\ell-1})$ is said to be* consistent *with strategy φ if $\forall 0 \leq i < \ell$: $s_i = \varphi(G_i)$.*

Closeness centrality amounts to an important index of social capital that captures a node's ease in accessing information, social support and other resources [12, 31, 1]. Thus we use closeness centrality here to indicate the position advantage of nodes. For any connected $G = (V, E)$ and $v \in V$, define

$$C_{\mathsf{Cls}}(v) = \frac{|V| - 1}{\sum_{u \in V \setminus \{v\}} \mathsf{dist}(u, v)}.$$

A higher value of $C_{\mathsf{Cls}}(v)$ implies that v is in general closer to other nodes, thus it occupies a better network position. The Cls-*rank* of v is the percentage of nodes whose closeness centrality are higher or equal to $C_{\mathsf{Cls}}(v)$:

$$\mathsf{rank}_{\mathsf{Cls}}(v) = |\{u \in V \mid C_{\mathsf{Cls}}(u) \geq C_{\mathsf{Cls}}(v)\}|/|V|.$$

We assume that the goal of v is to gain a higher closeness centrality (or a low $\mathsf{rank}_{\mathsf{Cls}}$). One way to achieve this is to build a tie between v and all nodes in the network. However, establishing new relationships requires time, efforts and resources. To identify realistic solutions, one needs to define *costs* of relationship building. Here we consider temporal and establishment costs. Temporal cost is the number of rounds in the NB process and coincides with the number of edges created for v. The *proximity principle* states that ties are generally more difficult to establish between nodes that are further apart (e.g. reciprocal of distance is a score for link prediction [24]). We thus define establishment cost as the sum of distance between v and its linked nodes (prior to edge creation).

Definition 4. *For an NB process $(G_0, G_1, \ldots, G_\ell, s_0, s_1, \ldots, s_{\ell-1})$,*

1. *the* temporal cost *is ℓ, and*
2. *the* establishment cost *is $\sum_{i=1}^{\ell-1} \sum_{u \in S_i} \mathsf{dist}_{G_i}(v, u)$.*

We are now ready to present the *dynamic network building (DNB) problem*: Given a connected social network G and newcomer v, the problem asks for an NB strategy φ such that any NB process consistent with φ will have high $C_{\mathsf{Cls}}(v)$ (or small $\mathsf{rank}_{\mathsf{Cls}}(v)$) value, and low temporal and and establishment costs. Fig. 1 displays a simple example where the graph evolves with the dynamic BA mechanism (see below); the newcomer gains a high centrality in three rounds.

The DNB problem differs from building relations in a static networks, which has been discussed in [25]: (1) As the network evolves, the NB process may last indefinitely where v tries to improves and maintains its centrality; (2) Network evolution forces v to balance between current knowledge with future predicted outcome. For example, linking to a central node will improve v's centrality quickly, but also incurs a high cost; on the other hand, linking to a low-centrality node may seem undesirable in the current network, but this link may improve the newcomer's centrality in the future. In this way, the evolution mechanism significantly impacts the newcomer's strategy.

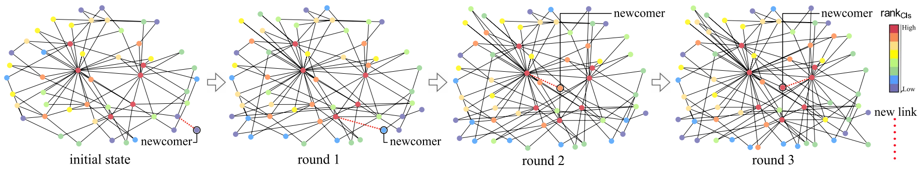

Fig. 1. A newcomer adds one edge at each round and achieves high centrality after 3 rounds, while the network is evolving.

3 Exploratory and Exploitative Strategies

Exploitative strategy. This strategy utilizes existing social proximity of the node v and searches for the most promising node that lies within a pre-defined distance from v. Fix as parameter a centrality index $C_*: V \to \mathbb{R}$ for nodes. Let $d \geq 2$ be a *proximity threshold*. When creating an edge, the strategy traverses through nodes with distance $\leq d$ from v, and picks a node found with maximum C_* value. Proc. 1 defines one round of the exploitative strategy. For the choice of C_*, we use standard centrality metrics that reflect aspects of social capital. The variety of centrality metrics below allow us to examine different potential heuristics, which may not always correlate [6].

1. *Degree*: $C_{\mathsf{Deg}}(u) = \{w \mid uw \in E\}$.
2. *Betweenness*: $C_{\mathsf{Btw}}(u) = \sum_{s \neq u \neq t \in V} |P_{st}(u)|/|P_{st}|$ where P_{st} is the set of shortest paths between s and t, $P_{st}(u) \subseteq P_{st}$ is those shortest paths that contain u.
3. *Closeness*: $C_{\mathsf{Cls}}(u)$.

We denote using Ld-Deg, Ld-Btw and Ld-Cls the local heuristics with centrality C_{Deg}, C_{Btw}, C_{Cls}, resp. When $d = 2$, the newcomer always links to a "friend-of-friends", a strategy studied in [20].

Procedure 1 Given $G = (V, E)$, $v \in V$

$U \leftarrow \{u \mid \mathsf{dist}(u, v) \leq d\}$
return a node $u \in U$ with maximum $C_*(u)$

Exploratory strategy. This strategy explores beyond the social proximity of v and links v with promising nodes in potentially distant parts of the network: The strategy takes a centrality index $C_*: V \to \mathbb{R}$ and a *distance threshold* $d \in \mathbb{N}$ as parameters. Call all nodes within distance d from v *covered*; at each round, the strategy will pick an uncovered node that has maximum C_*-value. Procedure 2 describes a single round of this strategy. We use Gγ-Deg, Gγ-Btw and Gγ-Cls to denote the exploratory heuristic that use $C_{\mathsf{Deg}}, C_{\mathsf{Btw}}$ and C_{Cls}, respectively.

Since the exploration creates edges to nodes which may be far away from v, by definition it bridges different parts of the network more quickly than the exploitation. Indeed, this can be verified using a simple example: Fix a large natural number n. Consider the *path graph* L_{2n+1} with $2n+1$ nodes (with nodes $v_1, v_2, \ldots, v_{2n+1}$ and edges $v_1v_2, v_2v_3, \ldots, v_{2n}v_{2n+1}$). Say v_1 is the newcomer. Assuming the network is static, G2-Cls builds $O(1)$ edges from v_1 (e.g. to $v_{n+1}, v_{n-2}, v_{n+2}$) and gives v_1 the highest closeness centrality. On the contrary, the exploitative strategy will create $\Omega(n)$ new edges to have the highest closeness centrality. In the next section, we compare the two strategies above through experiments on standard network evolution mechanisms and real-world data.

Procedure 2 Given $G = (V, E)$, $v \in V$

Find maximum $\gamma' \le \gamma$ with $V \neq \{u \mid \mathsf{dist}(v, u) \le \gamma'\}$
$U \leftarrow V \setminus \{u \in V \mid \mathsf{dist}(v, u) \le \gamma'\}$
return a node $u \in U$ with maximum $C_*(u)$

4 Contrasting Exploitative and Exploratory Strategies

4.1 Network Evolution Mechanisms

We consider three standard *network formation models*. Originally, each of these model were used to generated static networks. Here we extend them so that they entail mechanisms for network generation and evolution.

(a) Dynamic ER Model. The Erdös-Renyi (ER) random graph model adds edges between nodes as Bernoulli random variables with probability p. The degree distribution in the resulting graph thus follows a binomial distribution $B(n, p)$ [7]. We extend the model to a *death-birth* evolution model: Start from an ER random graph and introduce parameter $r \in [0, 1]$. At each round, first remove a randomly chosen set of nodes of size rn (i.e., death); then, add rn new nodes and link them with nodes in the graph with probability p (i.e., birth). It is clear that the operation preserves the binomial degree distribution $B(n, p)$.

(b) Dynamic BA Model. The Barabási-Albert model generates scale-free networks through a preferential attachment mechanism [3]. To define an evolving network model, we follow Barabási's dynamic extension. The key ideas include *growth*, *link establishment* and *node deletion* [2]. (a) Growth takes a rate $g \in [0, 1]$ and adds gn new nodes; adds m edges from each new node to an existing node with probability $k_i / \sum_{v_j \in V} k_j$ where k_j is the degree of v_j, $\forall v_j \in V$. (b) Link establishment takes a parameter $\lambda \in [0, 1]$; selects λn pairs of nodes in the current graph and randomly creates edges between these pairs; the probability of creating edge v_iv_j is proportional to k_ik_j. (c) Node deletion takes a rate $d \in [0, 1]$; picks dn nodes, and remove each of them (say, v_i) with probability $\frac{1/k_i}{\sum_j 1/k_j}$.

(c) Dynamic WS-model. The Watts-Strogatz model starts from a regular lattice, and performs random edge rewirings (with probability β) to obtain small-world networks, which have high levels of clustering and low average path length

[34]. We extend the process to an evolution mechanism: after initialization, the network evolves in each round by rewiring those previously-rewired edges to a random node. This dynamic network preserves the small-world property.

Table 1 summarizes the parameters used in our experiments. We choose these values either because they are standard choices used by others (e.g. $m = 2$ for dynamic BA [3]), or they ensure a gradual and smooth change at each round (e.g. for dynamic ER and WS models). '

Table 1. Parameters for evolving network models.

Evolving Model	Parameters
Dynamic ER	$r = 5\%$, $np = 4$
Dynamic BA	$m = 2, \lambda = 4\%, g = 5\%, d = 2\%$
Dynamic WS	$\beta = 0.2$

Experiment 1. (Costs) Through simulating DNB processes, we compare the temporal and establishment costs between the exploitative and exploratory s-trategies. DNB processes are generated by applying the heuristics in conjunction with the ER, BA and WS models. As a DNB process may have indefinite length, we need a termination condition to specify when the simulation stops. A natural method is to set a (high) threshold on centrality $C_{\mathsf{Cls}}(v)$, or set a (small) threshold on $\mathsf{rank}_{\mathsf{Cls}}(v)$, such that the process terminates once the threshold is met. There are problems with this approach: (1) It is difficult to determine a desired $C_{\mathsf{Cls}}(v)$ that facilitates fair comparisons across all evolving models. (2) In certain cases (e.g. WS model), closeness centrality of nodes are distributed within a small range; Hence, a node with low centrality may still have a small $\mathsf{rank}_{\mathsf{Cls}}$. These concerns motivate us to set a termination condition based on the ratio $g(v) = C_{\mathsf{Cls}}(v)/\mathsf{rank}_{\mathsf{Cls}}(v)$; we introduce a threshold ζ such that the process terminates at the first round when $g(v) \geq \zeta$ is satisfied.

We generate 10 networks of each size $n = 100, 200, 500, 1000$ using any network models above. We compare the exploitative with the exploratory strategies by running L2-$*$, L3-$*$ and G2-$*$ heuristic on each graph. Note that L2-$*$ only links v to her "friends-of-friends" which amounts to the most "local" exploitative strategy; L3-$*$ reaches beyond this local proximity and results in a different performance (see below); G2-$*$ is an exploratory strategy which tries its to link v to nodes outside of her local proximity. We do not include results for G3-$*$ as they are very similar to G2-$*$. For any generated graph G and heuristic, we do the following: (1) Build an edge between the newcomer v and a randomly chosen node in G.[3] (2) Apply the heuristic to the evolution mechanism (corresponding to the network formation model) to generate 100 DNB processes; the threshold

[3] We run the same experiment while linking v with initial nodes of different Cls-rank ($10\% - 90\%$). The resulting temporal and establishment costs are very similar. This shows that the rank of the initial node does not significantly affect the performance of the strategies.

ζ is 33. [4] (3) After the DNB process terminates, measure the resulting temporal and establishment costs. (4) Finally, record the average costs among all DNB processes of the same evolving network model, initial network size, and heuristic.

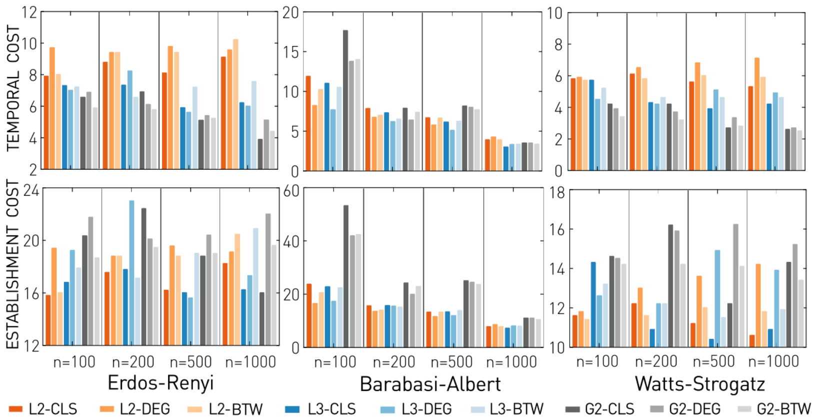

Fig. 2. Average temporal and establishment costs of simulated DNB processes by different heuristics on networks of different sizes.

A few facts stand out from the results in Fig. 2: **(i)** Temporal costs are mostly below 10, suggesting that a small number of edges are built by the strategies, even when the size of the network becomes 1000. **(ii)** For ER and WS, exploration results in a lower temporal cost compared to the exploitation. **(iii)** For the scale-free model BA, the number of edges built decreases as n increases. This may be due to the expanding nature of the dynamic BA model and the skewed degree distribution. As a result, exploitation creates less or similar numbers of edges than exploration. **(iv)** In general, exploration results in a much higher establishment costs. This is easy to understand: the strategy links to distant nodes from v. The L3 heuristics builds less edges than L2 due to the ability to traverse to a wider part of the graph. **(v)** The effects of the centrality metrics C_* vary with graph models: For ER, closeness centrality is in general preferred, while for BA, degree centrality is slightly more preferred. For WS, C_{Btw} is better for exploration, while C_{Cls} is better for exploitation as the graph becomes large.

We then plot v's Cls-rank as new edges are created during the DNB process; See Fig. 3. $\mathsf{rank}_{\mathsf{Cls}}(v)$ reaches $< 1\%$ after 10 rounds under all heuristics. G2 improves Cls-rank faster than L2 heuristics. This is most evident for L2-Deg and L2-Cls in WS-graphs: the high level of clustering may make it hard for

[4] The value $\zeta = 33$ reflects the fact that the desired $C_{\mathsf{Cls}}(v) \approx 1/3$ and $\mathsf{rank}_{\mathsf{Cls}}(v) \approx 1\%$. Other outcomes with $g(v) = 33$ that have either considerably lower centrality and Cls-rank, or considerably higher centrality and Cls-rank (e.g. $(C_{\mathsf{Cls}}(v), \mathsf{rank}_{\mathsf{Cls}}(v))$ is $(1/6, 0.5\%)$ or $(1, 3\%)$) have been empirically shown to be unlikely.

exploitation to get out of a dense cluster, but clustering does not seem to pose a problem for L2-Btw.

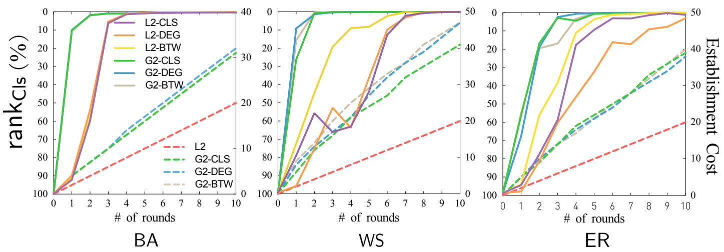

Fig. 3. Changes to $\mathsf{rank}_{\mathsf{Cls}}(v)$ (solid lines) and the establishment costs (dashed lines) during 10 rounds of the DNB process. The network starts with 1000 nodes and v initially connects to a node with lowest centrality.

Experiment 2. (Embeddedness and clustering) Exp. 1 implies that, to some extend, exploitation and exploration perform on par with each other from a centrality perspective. In building relationship, trust, tie strength and role integrity are other important dimensions not captured by centrality alone [16]. These notions are closely affected by two concepts:

1. *Embeddedness* in a social network refers to the degree to which an individual is constrained by social relationships and is often viewed as a platform for trust [17]. The embeddedness of an edge between x, y is defined as the size of their shared neighborhoods $D(x) \cap D(y)$ [11]. We define $\mathsf{embed}(v)$ as a normalized sum of embeddedness: $\mathsf{embed}(v) = \frac{\sum_{vu\in E} |\{w\in V | wu, wv\in E\}|}{(|V|-1)(|V|-2)}$.
2. *Clustering coefficient* of a node measures the probability of two randomly chosen friends of the node are also friends and relates to the self-identify of individuals [4]: $\mathsf{cc}(v) = \deg(v)(\deg(v)-1)/2$.

We measure $\mathsf{embed}(v)$ and $\mathsf{cc}(v)$ during DNB processes; See Fig. 4. Remarkably, $\mathsf{embed}(v)$ increases drastically with exploitation, while staying close to 0 with exploration. The clustering coefficient $\mathsf{cc}(v)$ also stays close to 0 with exploration, while with exploitation, it quickly rises to a very high level in the first few rounds, and drops down to below 0.1 after 10-15 rounds. This highlights the newcomer's ability to cut across different clusters. Overall, this experiment demonstrates the crucial difference between the strategies: while both strategies improve v's closeness centrality, exploitation enables a higher embeddedness and clustering coefficient which positively correlates with tie strength and trusts on its social relations.

4.2 Real-World Evolving Networks

We next take real-world evolving network data as case studies.

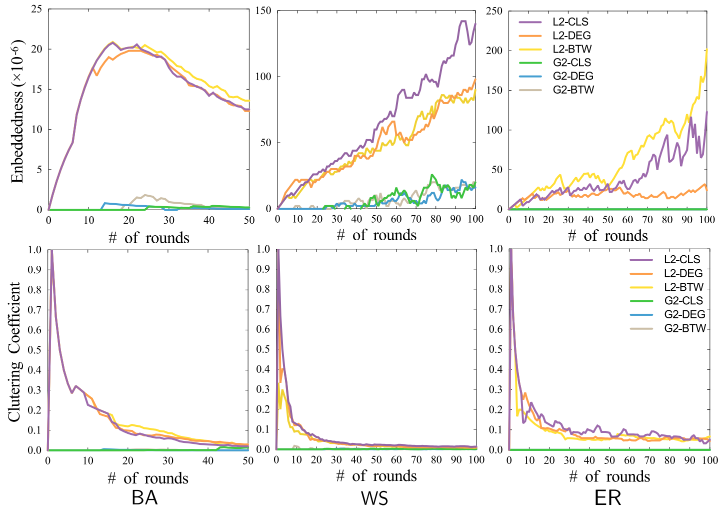

Fig. 4. Changes to v's embeddedness and cc as edges are added to v. The network starts with 1000 nodes where v is connected to a node with the lowest centrality.

Trade network 1949–2009. The first case study utilizes annual bilateral trade and GDP data between 1949 and 2009 [14]. Using the same approach as in [21], we measure the level of trade between two countries based on exchanged goods divided by GDP[5]: An edge represents *trade partnership*, which exists between two countries a and b if the export of one to the other is $\geq 0.5\%$ of its GDP in that year. International trade has expanded enormously in the last 60 years. The number of countries in the data set grows from 38 (with 43 partnerships) in 1949 to 176 (with 1229 partnerships) in 2009. We perform a thought experiment: Suppose the year is 1949 and v is a newly established economic zone. Adding one trade partner every two years, what countries should v establish trade with in the next 60 years to gain a central position in the world market?

We start by linking v with New Zealand[6], and apply L2 and G2 to build edges biennially. Fig. 5 plots the resulting Cls-rank, embeddedness and clustering coefficient of v through time. All heuristics, with the exception of G2-Cls, quickly bring $C_{\mathsf{Cls}}(v)$ to within top 10% of all countries. The ranking then fluctuates, which may be explained by consistent expansions of the network. On the other hand, L2 heuristics lead to considerably higher embeddedness and clustering coefficient than G2, suggesting higher integrity of v's trade partnerships. As Table 2 shows, the list of partner countries produced by L2-Deg is remarkably consistent with real historical events, connecting to countries during major boom: e.g. USA & UK were the leading powers in the 1950s, West Germany's Wirtschaftswunder

[5] Measured in terms of Parity Purchase Power of USD in 2005

[6] We choose New Zealand as the initial node connected to v. This is because New Zealand has low-mid level Cls-rank.

(1950s), Japan's miracle (1960s), Germany reunification (1990), China's joining of WTO (early 2000s). On the contrary, the list by G2-Deg is somewhat surprising, comprising countries during economic or political crisis such as Argentina (1953), Lebanon (1979), GDR (1981), Burundi (1991) and Zimbabwe (2007). This reflects numerous discussions on embeddedness in economical theories [17]: exploitation tends to produce meaningful alliance and cooperation, while exploration tends to be more opportunistic and speculative. A video visualizing the DNB process computed by L2-Deg is in `https://youtu.be/sz5BGHtwPu4`

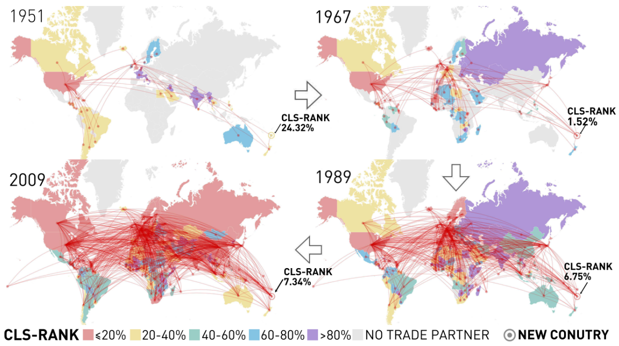

Fig. 5. Changes to Cls-rank, embed(v), cc(v) resulted from all heuristics.

Table 2. Lists of countries computed by L2-Deg and G2-Deg, respectively.

Year	L2-Deg	G2-Deg	Year	L2-Deg	G2-Deg
1951	USA	Saudi Arabia	1981	Yemen Arab	East Germany
1953	UK	Argentina	1983	Singapore	Dominica
1955	EI Salvador	USA	1985	S. Korea	Turkey
1957	Finland	USSR	1987	Switzerland	India
1959	West Germany	Netherland	1989	Sweden	Djibouti
1961	Netherland	France	1991	Germany	Burundi
1963	Kuwait	Japan	1993	Thailand	Pakistan
1965	France	Italy	1995	Russia	Czech
1967	Japan	Denmark	1997	Canada	Cyprus
1969	Liberia	Singapore	1999	China	Kenya
1971	Libya	Norway	2001	Denmark	Togo
1973	Italy	Spain	2003	Austria	Turkmenistan
1975	UAE	Portugal	2005	Poland	Zambia
1977	Belgium	Brazil	2007	India	Zimbabwe
1979	Spain	Lebanon	2009	Malaysia	Cuba

Contact networks. We then take two evolving physical contact networks: The first records face-to-face contacts among ∼110 attendees of ACM Hypertext 2009 conference during 2.5 days [18]. The second records contacts among ∼100 employees in a French workplace June 24 to July 3, 2013 [13]. In both data sets, contacts are updated every 20 seconds. We ask the question: If a newcomer attends the conference, or joins the workplace, how does she utilize face-to-face contacts to reach a central position of the network? We simulate DNB processes on the *accumulated network*, i.e., networks constructed by accumulating edges in previous times, using the $\mathsf{G2}$ and $\mathsf{L2}$ heuristics. For the first data set, an edge is built every 15 minutes, and for the second data set, an edge is built very hour. Fig. 6 plots the changes on the newcomer's Cls-rank. All heuristics have similar results to improve the Cls-rank as time progresses. As the network is small and edges are accumulated, the heuristics somehow fail to improve newcomer's rank in the last three days of the second network beyond 20%. The exploitative strategy, however, gives much smaller establishment costs.

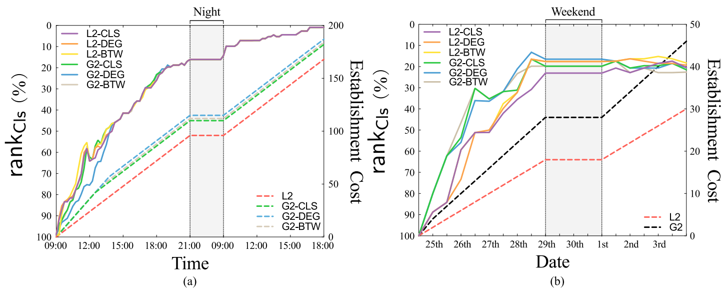

Fig. 6. (left) ACM Hypertext 2009 contact network; (right) French workplace contact network.

5 Combining Exploitation and Exploration

We use UCB1 (Upper Confidence Bound) algorithm to combine exploitative and exploratory strategies. Exploitative and exploratory strategies are regarded as two choices and in each round of DNB problem and one strategy is selected in each round which achieves the maximum value:

$$UCB(i) = \underbrace{\frac{\sum_{j=1}^{N} \mathsf{rank}_{\mathsf{Cls}}(r_{ij}-1) - \mathsf{rank}_{\mathsf{Cls}}(r_{ij})}{N}}_{average\, function} + \underbrace{\sqrt{\frac{2\ln t}{n_i}}}_{padding\, function}$$

where t is the number of rounds passed, n_i is the times that strategy i is selected, r_{ij} is the $j-$th round that strategy i is selected, N is the total number of times when strategy i is selected. $Average\, function$ denotes average benefit that the newcomer has got so far and we use the difference value of newcomer's $\mathsf{rank}_{\mathsf{Cls}}$ between two contiguous rounds to define the benefit that the newcomer gets after making the decision in each round. $Padding\, function$ denotes an approximation of the uncertainty on strategy i, the more times that strategy i is selected, the less uncertainty it has.

Intuitively, we can get from the this combining method that if one strategy seems to have no capacity to increase newcomer's $\mathsf{rank}_{\mathsf{Cls}}$, try the other one.

In the beginning of DNB process, we use exploratory and exploitative s-trategies in round1 and round2, resp. Then, in the following rounds, select the strategy achieving the bigger value in UCB. Procedure 3 shows the mechanism of DNB−UCB in a DNB process in detail, where v is the newcomer.

Procedure 3 DNB-UCBγ-$*$: Given $G = (V, E), v \in V$

Initialization: Select $\mathsf{L}\gamma$-$*$ and $\mathsf{G}\gamma$-$*$ at round 0 and round 1 resp.
if $UCB(\mathsf{L}\gamma\text{-}*) \geq UCB(\mathsf{G}\gamma\text{-}*)$ **then**
  Find maximum $\gamma^{'} \leq \gamma$ with $V \neq \{u \in V | \mathsf{dist}(v, u) \leq \gamma^{'}\}$
  $U \leftarrow V \backslash \{u \in V | \mathsf{dist}(v, u) \leq \gamma^{'}\}$
  return a node $u \in U$ with maximum $C_*(u)$
else
  $U \leftarrow \{u \in V | \mathsf{dist}(u, v) \leq \gamma\}$
  return a node $u \in U$ with maximum $C_*(u)$
endif

6 Conclusion and Future Work

The paper proposes dynamic network building problem and concentrates on contrasting exploratory and exploitative strategies. While both strategies lead to high closeness centrality of the newcomer, the local strategy tends to have lower establishment cost and resembles natural network building process. More experiments are needed to reveal further insights on these strategies. The general question is, How much social capital does each strategy bring to the newcomer? The answer of this question requires precise definitions of social capital beyond centrality and embeddedness [10]. Furthermore, [20] describes a network formation model where “meeting strangers” is combined with “friends-of-friend”. Following the same spirit, one asks: How to amalgamate local and global strategies to achieve better network building? As the local strategy does not necessarily require global information about the network, a third question asks: can we utilize the local strategy to design distributed algorithms for efficient network building in very large social network? We will explore these questions as future works.

References

1. F. Agneessens, S. P. Borgatti, and M. G. Everett. Geodesic based centrality: Unifying the local and the global. *Social Networks*, 49:12–26, 2017.
2. A.-L. Barabási. *Network Science*. Cambridge University Press, 2016.

3. A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
4. P. S. Bearman and J. Moody. Suicide and friendships among american adolescents. *American Journal of Public Health*, 94(1):89–95, 2004.
5. M. Benner and M. Tushman. Exploitation, exploration, and process management: The productivity dilemma revisited. *Acad. Manage. Rev.*, 28(2):238–256, 2003.
6. J. Bolland. Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. *Social Networks*, 10:233–253, 1988.
7. B. Bollobás. *Random Graphs (2nd ed.)*. Cambridge University Press, 2001.
8. P. Bonacich. Power and centrality: A family of measures. *Am. J. Sociol*, 92:1170–1182, 1987.
9. S. Borgatti. Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory*, 12(1):21–34, 2006.
10. R. Burt. *Brokerage and closure: An introduction to social capital*. Oxford University Press, 2005.
11. D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.
12. L. Freeman. Centralityinsocialnetworks:conceptualclarification. *SocialNetworks*, 1:215–239, 1979.
13. M. Génois, C. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, and A. Barrat. Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers. *Network Science*, 3(3):326–347, 2015.
14. K. S. Gleditsch. Expanded trade and gdp data. *Journal of Conflict Resolution*, 46(5):712–724, 2002.
15. M. Granovetter. The strength of weak ties. *Am. J. Sociol*, 78(6):1360–1380, 1973.
16. M. Granovetter. Economic action and social structure: the problem of embeddedness. *Am. J. Sociol*, 91(3):481–510, 1985.
17. R. Gulati. Alliances and networks. *Strategic Manage. J.*, 19:293–317, 1998.
18. L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J.-F. Pinton, and W. V. den Broeck. Whats in a crowd? analysis of face-to-face behavioral networks. *Journal of Theoretical Biology*, 271:166–180, 2011.
19. M. Jackson. *A Survey of Models of Network Formation: Stability and Efficiency, Group Formation in Economics*, pages 11–57. Cambridge University Press, 2005.
20. M. Jackson and B. Rogers. Meeting strangers and friends of friends: How random are social networks? *Am. Econ. Rev*, 97(3):890–915, 2007.
21. M. O. Jackson and S. Nei. Networks of military alliances, wars, and international trade. *Proc. Nat. Acad. Sci. USA*, 112(50):15277–15284, 2015.
22. S. H. Lee, J. Cotte, and T. J. Noseworthy. The role of network centrality in the flow of consumer influence. *Journal of Consumer Psychology*, 20(1):66–77, 2010.
23. J. Liu and A. Moskvina. Hierarchies, ties and power in organizational networks: model and analysis. *Social Netw. Analys. Mining*, 6(1):106:1–106:26, 2016.
24. L. Lü and T. Zhou. Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and its Applications*, 390(6):1150–1170, 2011.
25. A. Moskvina and J. Liu. How to build your network? a structural analysis. In *Proc. of IJCAI-16*, pages 2597–2603. AAAI, 2016.
26. M. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences USA*, 98:404–409, 2001.
27. M. Newman. *Networks: An Introduction.* Oxford, UK: Oxford University Press, 2010.
28. T. Osugi, K. Deng, and S. Scott. Balancing exploration and exploitation: A new algorithm for active machine learning. In *Proc. ICDM 2005*, pages 330 –337, 2005.

29. J. F. Padgett and C. K. Ansell. Robust action and the rise of the medici, 1400-1434. *Am. J. Sociol*, 98(6):1259–1319, 1993.
30. F. R. Pitts. The medieval river trade network of russia revisited. *Social Networks*, 1(3):285–292, 1978.
31. R. T. Sparrowe, R. C.Liden, S. J.Wayne, and M. L. Kraimer. Socialnetworksandtheperformanceof individualsandgroups. *AcademyofManagementJournal*, 44:316–325, 2001.
32. B. Uzzi and S. Dunlap. How to build your network. *Harvard Business Review*, 83(12):53–60, 2005.
33. J. Vermorel and M. Mohri. Multi-armed bandit algorithms and empirical evaluation. In *Proc. ECML 2005*, pages 437–448. Springer, 2005.
34. D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.