Gerlach and Blumenthal
Applied Network Science (2023) 8:24
https://doi.org/10.1007/s41109-023-00550-x

Applied Network Science

RESEARCH

Open Access

On the role of network topology in German-Jewish recommendation letter networks in the early twentieth century

Lisa Gerlach[1] and David B. Blumenthal[2*]

*Correspondence: david.b.blumenthal@fau.de

[1] Faculty of History, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany
[2] Biomedical Network Science Lab, Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Werner-von-Siemens-Str. 61, 91052 Erlangen, Germany

Abstract

Recommendation letters were an important instrument for orchestrating Jewish emigration from Germany in the late nineteenth and early twentieth century. Here, we present network-based analyses of manually collected meta-data from recommendation letters targeted at the Hebrew University (HU) in Jerusalem. Using standard semi-supervised node classification techniques and differential node centrality analyses, we show that the position of a recommendation letter in content-agnostic recommendation network models is predictive of its success, i.e., of whether or not the letter led to the recommendee obtaining a position at the HU. In particular, we show that authors of successful recommendation letters assume more central positions within the networks than authors of unsuccessful letters, while the opposite holds for the recommendation letters' receivers. Beyond our application, these results showcase the potential of using network models for generating historical insights. Both the letter meta-data records and Python code to reproduce our analyses are available on GitHub: https://github.com/bionetslab/cornet.

Keywords: Recommendation letter networks, Jewish emigration from Nazi Germany, Hebrew University

Introduction

Recommendation letters can still make or break an academic career. Within the early days of the Hebrew University (HU) in Jerusalem, recommendation letters were an even more powerful tool in helping scholars to emigrate out of Nazi Germany. The HU was officially founded in 1918 and started operating in 1925 (Bentwich 1953, 1961). As the intellectual flagship of the zionist movement, the HU was supposed to become a haven in a world full of chaos and danger, in which Jews would find shelter in the darkest times (Dauben and Robinson 1995)—to paraphrase Chaim Weizmann's opening speech. During the early days of the young institution, positions were scarce but so was the interest of Western European Jewish academics to permanently migrate to Palestine. In fact, in the early days of the HU, members of the HU were reaching out to former colleagues to advertize open positions in Jerusalem. With little budget and highly ideological

Springer Open

© The Author(s) 2023. **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by/4.0/.

approaches, the HU aimed to further a better understanding of their surroundings and of Jewish religion, language, and culture (Gerlach 2022).

This dramatically changed when the so-called "Law for the Restoration of the Professional Civil Service" (*Gesetz zur Wiederherstellung des Berufsbeamtentums*) of April 17, 1933, forced many German-Jewish academics out of their positions at German universities. The HU was eager to help their German colleagues, many of them well known and well respected in their fields. But the young institution had to face the reality of financial boundaries. They had more applications than open positions (Bentwich 1961; Erel 1983). The University started to gather funding, establish new positions, and fill them primarily with people needing to leave Germany and later Western Europe. All whilst trying to build a preferably undiluted Zionist profile and recruit experts needed to fill in desired research focuses of the institution (Erel 1983).

Recommendation letters were part of the application documents scholars would send in. Many recommendation letters looked similar regarding their general structure and many of the candidates were equally qualified (Gerlach 2022). In view of this, we investigated if the success of a recommendation letter (i.e., whether or not the candidate obtained a position at the HU) did not only depend on its content but also on its position in the topology of the correspondence network induced by the recommendation letters. To answer this question, we constructed content-agnostic networks from the meta-data of the recommendation letters, which we analyzed via standard semi-supervised node classification techniques and differential node centrality analyses. Our analyses clearly show that success information is indeed encoded in our content-agnostic networks. In particular, we observe that senders of successful letters tend to have larger node centralities than senders of unsuccessful ones. For the recommendation letters' receivers, we make the opposite observation, potentially indicating that persons targeted by a large number of letters could dedicate less resources to each individual application.

In addition to our analyses, we present a manually generated dataset containing meta-data records for 903 recommendation letters. 298 records describe letters targeting the HU; the remaining letters target three other potential host institutions for Jewish persons who tried to emigrate from Germany (see details below). These letters were not used for our analyses, since we have access to success information only for the letters targeting the HU. Finally, we provide a well-documented Python package which not only allows to fully reproduce all reported results but can also be used for analyzing other historical correspondence letter networks. When developing the package, we aimed at maximum possible simplicity, such that it is usable also for historians with little programming experience. Both the dataset and the Python package are publicly available on GitHub: https://github.com/bionetslab/cornet.

Materials and methods

Manual generation of meta-data records for 903 recommendation letters

Recommendation letters targeting the HU were mainly sourced in the archive of the HU in Jerusalem with additions kept by the Central Zionist Archive (Jerusalem), the National Library of Israel (Jerusalem), and the Yad Weizmann Archive (Rehovot). Successful letters which led to the recommendee obtaining a position at the HU can be found in the personal files of German-Jewish scholars who, at some point between

1918 and 1945, held a position at the HU. Unsuccessful letters can be found and in dedicated files where unsuccessful applications had been archived. In addition to records targeting the HU, our dataset contains records for letters targeting three other potential host institutions (without success information):

- The Bleichröder Bank (Stern 2000), with records for letters archived at the Baker Library Special Collection, Harvard Business School (Boston) and the Archives of the Bank Rothschild Frères (Paris).
- The Arnhold Bank (Lässig 1997), with records for letters archived at the Sächsisches Staatsarchiv (Dresden), the Nietzsche Haus (Sils), the Siemens Historical Institute (Berlin), and the Landesarchiv Baden-Württemberg (Stuttgart).
- The American Friends Service Committee (Austin 2012), with records for letters archived at the United States Holocaust Memorial Museums Archives (Washington, D.C.).

Each letter was manually annotated as "Formal Recommendation", "Informal Recommendation", "Request for Recommendation", or "Introduction". We classified a letter as a formal recommendation if it contains the following five parts: First a greeting, stating the context of the letter. Second, a comment about how and how well the sender and the recommendee knew each other. Third, a lengthy description of the recommendee. Fourth, the actual recommendation—often an assessment on the recommendee's suitability for a position at the targeted institution. And fifth, a formal closing. Informal recommendations are recommendation letters which do not follow this strict structure; e.g., the recommendation might be part of a letter of mixed content which also addressed non-recommendation related issues. Request for recommendations are letters in which the hiring institution asks trusted experts in the field of interest to write a recommendation letter for a potential candidate, which either has already been sought out by the institution and needs further confirmation or has to be chosen by the recommending expert. Introductions are letters which only introduce a recommendee at the targeted host institution by name and often announce an upcoming visit or an interest in a collaboration.

From the compendium of recommendation letters described above, we manually generated meta-data with the following attributes for each letter:

- "Sender": The sender of the recommendation letter, acting as recommender.
- "Receiver": The person to whom the recommendation letter was addressed.
- "Written for": The recommendee for whom the letter was written.
- "Date": The letter's date.
- "Institution": The institution where the recommendee tried to obtain a position.
- "Success": A binary flag indicating whether the recommendee obtained a position (only for letters targeting the HU).
- "Recommendation Type": The recommendation letter type as described above. Note that we did not use this attribute for the analyses reported in this article.

In total, our dataset contains meta-data records for 903 recommendation letters. Table 1 provides an overview of the how the records are distributed over the four targeted host

Table 1 Distribution of host institutions targeted by the recommendation letters and counts of successful and unsuccessful letters

Institution	# Letters	# Successful	# Unsuccessful
Hebrew University	298	150	148
Bleichröder Bank	61	N/A	N/A
Arnhold Bank	39	N/A	N/A
American Friends Service Committee	505	N/A	N/A

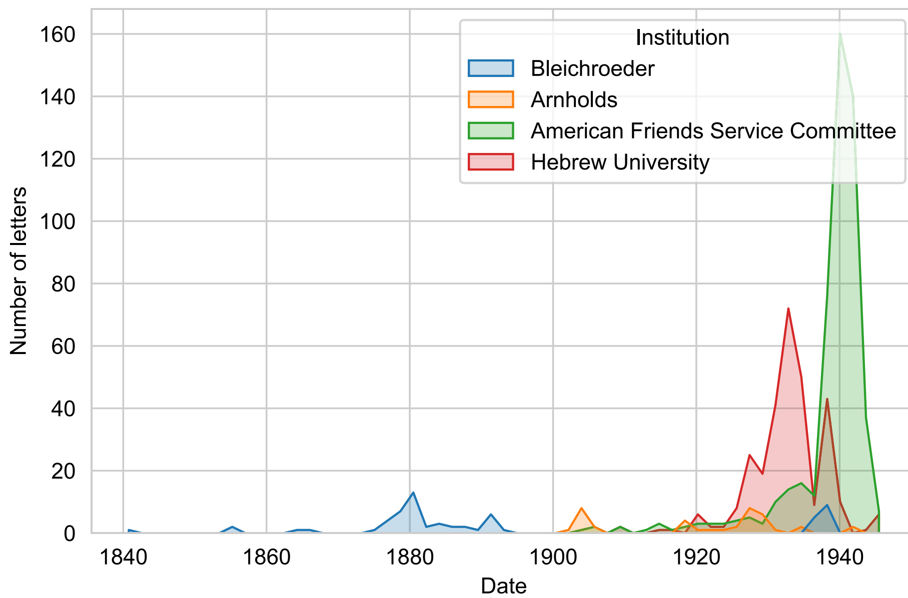

Fig. 1 Date distribution of recommendation letters split by targeted institution

Table 2 Numbers of nodes, edges, weakly connected components (WCCs), and nodes contained in largest WCC in networks constructed from recommendation letters targeting the HU

Network	# Nodes	# Edges	# WCCs	# Nodes in largest WCC
$\mathcal{M}$	242	298	24	163
$\mathcal{L}$	298	3117	66	223
$\mathcal{L}'$	298	3054	66	223
$\mathcal{D}$	242	250	24	163

institutions and Fig. 1 shows the date distributions for the different institutions. Note that only the records for letters targeting the HU were used for our network analyses.

Network construction

Let L be the set of all 298 meta-data records of letters targeting the HU. For each record $\ell \in L$, let $s(\ell)$ and $r(\ell)$ denote the values of its "Sender" and "Receiver" attributes, respectively. Moreover, let $success(\ell) \in \{\text{"True"}, \text{"False"}\}$ denote the value of ℓ's "Success" attribute. Furthermore, let $S = \{s(\ell) \mid \ell \in L\}$ and $R = \{r(\ell) \mid \ell \in L\}$ be the sets of all senders and receivers (note that $S \cap R \neq \emptyset$), and $P = S \cup R$ be the set of all persons acting as senders or receivers. We constructed five different networks from the meta-data records (see Table 2 for summary statistics):

- A directed multi-graph $\mathcal{M} = (P, L, \mathit{success})$, where a letter record $\ell \in L$ is interpreted as a directed edge from $s(\ell)$ to $r(\ell)$ and the success status is modeled as an edge attribute. $\mathcal{M}$ is a directed multi-graph because there are some sender-receiver pairs connected by multiple letter records, often with different values for the success attribute (typically, series of letters where all but the last were unsuccessful).
- The directed line graph $\mathcal{L} = (L, E_\ell, \mathit{success})$ obtained from $\mathcal{M}$, i.e., $(\ell, \ell') \in E_{\mathcal{L}}$ if and only if $r(\ell) = s(\ell')$. Note that $\mathcal{L}$ is not a multi-graph but a simple directed graph and that, in $\mathcal{L}$, the success status is a node attribute.
- The undirected version $\mathcal{L}' = (L, E'_{\mathcal{L}}, \mathit{success})$ of $\mathcal{L}$, where directed edges $(\ell, \ell') \in E_{\mathcal{L}}$ are replaced by their undirected counterparts $\ell\ell'$.
- A simple edge-weighted directed graph $\mathcal{D} = (P, E_{\mathcal{D}}, w)$ obtained from $\mathcal{M}$ by replacing multi-edges between a sender-receiver pair $(u,\ v)$ by a single directed edge weighted with its multiplicity $w(u, v) = |\{\ell \in L \mid s(\ell) = u \wedge r(\ell) = v\}|$ in $\mathcal{M}$. Note that, due to the aggregation of parallel edges with differing values of the success attribute, success is not modeled in $\mathcal{D}$.
- An inverted variant $\overleftarrow{\mathcal{D}} = (P, \overleftarrow{E_{\mathcal{D}}}, \overleftarrow{w})$ of $\mathcal{D}$, where $(v, u) \in \overleftarrow{E_{\mathcal{D}}}$ if and only if $(u, v) \in E_{\mathcal{D}}$ and $\overleftarrow{w}(v, u) = w(u, v)$ for each $(v, u) \in \overleftarrow{E_{\mathcal{D}}}$.

Due to the existence of sender-receiver pairs connected by multiple successful and unsuccessful letter records, the directed multi-graph $\mathcal{M}$ is the most natural network representation of our data. The other network representations should be considered as views of $\mathcal{M}$ that allow to carry out the analyses detailed below.

Semi-supervised prediction of success status

The main question we are addressing in this article is whether the topologies of the recommendation letter networks described above contain information about the success of records targeting the HU. To this end, we used the classical semi-supervised node classification methods by Zhu et al. (2003) and Zhou et al. (2003). These methods use label propagation techniques to complete node labelings of partially node-labeled undirected graphs. For our study, we ran these methods on the undirected line graph $\mathcal{L}'$. $\mathcal{L}'$ fits the input requirements of the semi-supervised node classifiers because, in $\mathcal{L}'$, the values of the "Success" attribute correspond to node labels.

We used a 5-fold cross-validation (CV) scheme, where we removed the ground-truth success labels in the test folds and computed predicted labels by propagating the ground-truth labels from the training folds. Mean CV accuracy was then computed as the mean fraction of correctly predicted success labels across the five test folds. To ensure robustness of the results, we repeated this protocol 100 times, leading to a sample of 100 mean CV accuracy values for each of the two semi-supervised methods. Note that we did not aim at constructing success classifiers that generalize to unseen recommendation letter records. Rather, our aim was to investigate if the topologies of the content-agnostic network models constructed from our specific dataset indeed contain information about the recommendations letters' success. This answer can be answered positively if the mean CV accuracy values exceed the accuracy of the baseline predictor which always predicts the majority class (in our case $150/297 \approx 0.51$, see Table 1).

To additionally test if the semi-supervised node classifiers merely pick up on random effects, we compared the obtained mean CV accuracies to mean CV accuracies obtained upon randomization of (1) the success labels and (2) the network topology of the undirected line graph $\mathcal{L}'$. If $\mathcal{L}'$ indeed contains information about success, the mean CV accuracies obtained for the real data should be significantly larger than the mean CV accuracies obtained upon randomization. For success label randomization, we shuffled the success labels on the nodes of $\mathcal{L}'$ prior to each of the 100 CV runs. For network randomization, before each CV run, we replaced the edges in $\mathcal{L}'$ by randomly sampled edges such that the node degrees are approximately preserved in expectation. This can be done by sampling edges $\ell\ell'$ between nodes $\ell, \ell' \in L$ with probability $p_{\ell\ell'} = (\deg(\ell) \cdot \deg \ell')/(2 \cdot |E'_{\mathcal{L}}|)$, where $\deg(\ell)$ and $\deg(\ell')$ are the node degrees of ℓ and ℓ' in $\mathcal{L}'$ (Chung and Lu 2002; Miller and Hagberg 2011). With this sampling strategy, the expected degree of node ℓ in the randomized network equals $\deg(\ell) \cdot (1 - \deg(\ell)/(2 \cdot |E'_{\mathcal{L}}|)) \approx \deg(\ell)$.

Differential node centrality analyses

We further analyzed if there are significant differences between node centrality distributions of, respectively, sender and receiver nodes in the weighted directed graph $\mathcal{D} = (P, E_{\mathcal{D}}, w)$ incident with edges corresponding to successful and unsuccessful letters. To this end, we computed the following widely used centrality measures for each node $u \in P$:

- The in-degree centrality, defined as

$$D(u) = \frac{\deg^-(u)}{n-1}, \tag{1}$$

  where $\deg^-(u)$ denotes the number of node u's incoming edges and $n = |P|$ is the number of nodes in $\mathcal{D}$.
- The closeness centrality (Bavelas 1950) in the version by Wasserman and Faust (1994) for graphs with several connected components:

$$CC(u) = \frac{|R(u)|}{n-1} \cdot \frac{|R(u)|}{\sum_{v \in R(u)} \mathrm{dist}\,(v,u)} \tag{2}$$

  Here, $\mathrm{dist}\,(v,u)$ denotes the shortest-path-distance from v to u in $\mathcal{D}$ and $R(u)$ is the set of all nodes $v \in P \backslash \{u\}$ such that u is reachable from v. Note that Bavelas' classic definition $CC(u) = (\sum_{v \in P \backslash \{u\}} \mathrm{dist}\,(v,u))^{-1}$ is meaningless for networks with several connected components, because it yields $CC(u) = 1/\infty = 0$ for all nodes.
- The harmonic centrality (Marchiori and Latora 2000), defined as

$$H(u) = \sum_{v \in P \setminus \{u\}} \frac{1}{\mathrm{dist}\,(v,u)}. \tag{3}$$

The harmonic centrality and the closeness centrality are closely related. The main difference is that the closeness centrality uses the weighted inverse of the arithmetic mean over the shortest-path-distances, while the harmonic centrality uses the inverse of the harmonic mean.

- The PageRank centrality (Brin and Page 1998), defined recursively as

$$PR(u) = \frac{1-d}{n} + d \cdot \sum_{v \in N^-(u)} \frac{w(v,u)}{W^+(v)} \cdot PR(v), \tag{4}$$

where $N^-(u)$ is the set of in-neighbors of node u, $w(v, u)$ is the weight of the edge (u, v) (i.e., its multiplicity in $\mathcal{M}$), $W^+(v)$ is sum of weights of edges outgoing from v, and $d \in (0, 1)$ is a hyper-parameter (set to the NetworkX default $d = 0.85$ for our study).

With respect to all four centrality measures, a node u is important if it has a high in-connectivity, or in other words, if it is an important target node. To also quantify out-connectivity (or importance-as-source-node), we additionally computed $D(u)$, $CC(u)$, $H(u)$, and $PR(u)$ in the inverted counterpart $\overleftarrow{\mathcal{D}}$ of $\mathcal{D}$ (note that, for each node $u \in P$, the in-degree centrality $D(u)$ in $\overleftarrow{\mathcal{D}}$ is equivalent to its out-degree centrality in the original graph $\mathcal{D}$). In the following, the resulting centrality measures are denoted by $\overleftarrow{D}(u)$, $\overleftarrow{CC}(u)$, $\overleftarrow{H}(u)$, and $\overleftarrow{PR}(u)$, respectively.

To test for differences in centrality between nodes incident with successful and unsuccessful letters, we constructed (pairwise non-exclusive) sets of successful and unsuccessful sender and receiver nodes:

$$S_{\text{success}} = \{u \in P \mid \exists \ell \in L : s(l) = u \wedge success(l) = \text{“True”}\} \tag{5}$$

$$S_{\text{no success}} = \{u \in P \mid \exists \ell \in L : s(l) = u \wedge success(l) = \text{“False”}\} \tag{6}$$

$$R_{\text{success}} = \{u \in P \mid \exists \ell \in L : r(l) = u \wedge success(l) = \text{“True”}\} \tag{7}$$

$$R_{\text{no success}} = \{u \in P \mid \exists \ell \in L : r(l) = u \wedge success(l) = \text{“False”}\} \tag{8}$$

Now, for each centrality measure $C \in \{D, CC, H, PR, \overleftarrow{D}, \overleftarrow{CC}, \overleftarrow{H}, \overleftarrow{PR}\}$ and each $X \in \{S, R\}$, we used the two-sided Mann–Whitney-Wilcoxon (MWW) test to assess whether there are significant differences between the centralities $\{\{C(u) \mid u \in X_{\text{success}}\}\}$ of successful sender or receiver nodes and the centralities $\{\{C(u) \mid u \in X_{\text{no success}}\}\}$ of unsuccessful ones. In total, we hence carried out sixteen MWW tests (eight centrality measures, separate tests for sender and receiver nodes).

Another widely used node centrality measure is the betweenness centrality (Freeman 1977), defined as $BC(u) = \sum_{v \neq u \neq w} \sigma_{v,w}(u)/\sigma_{v,w}$ with $\sigma_{v,w}$ the number of shortest paths from v to w and $\sigma_{v,w}(u)$ the number of shortest v-w-paths that pass through u. However, the betweenness centrality does not allow to distinguish between importance-as-target-node and importance-as-source-node because $\sigma_{v,w}(u)$ and $\sigma_{v,w}$ are invariant w.r.t. inversion of $\mathcal{D}$. Since this distinction is crucial for our analyses, we did not use the betweenness centrality for this study.

Implementation and justification for study design

We used the NetworkX (Hagberg et al. 2008) implementations of the semi-supervised node classifiers (with all hyper-parameters set to defaults), as well as of the network randomization with approximately preserved expected node degrees. Similarly, all node centrality measures were implemented with NetworkX, again using default hyper-parameters. MWW tests were implemented using SciPy (Virtanen et al. 2020). The entire analysis pipeline is provided in the Python package CorrNet (https://github.com/bionetslab/corrnet), which can also be used to analyze user-provided historical correspondence networks.

Note that we deliberately choose methods for which reference implementations are available in extremely popular and easily usable Python packages and that we consciously refrained from hyper-parameter tuning. There are two reasons for these choices. The first reason is that, as already mentioned above, the aim of this study is *not* to design a general-purpose success predictor for recommendation letter networks but rather to assess whether, *for our specific dataset*, success is encoded in content-agnostic meta-data networks. To answer this question *about the data*, using out-of-the-box methods rather than tailored approaches is preferable, because it decreases the risks of overfitting and confirmation bias. Moreover, a secondary objective of this study is to showcase how network science techniques can be used to generate historical insights in a way which is accessible to historians with some interest in the digital sciences but no strong algorithmic background (Althage et al. 2022). Using a simple study design is beneficial also for this purpose.

Results

Network topology is predictive of success

Figure 2 shows the distributions of mean CV accuracies obtained by the harmonic function classifier by Zhu et al. (2003) and local and global consistency classifier by Zhou et al. (2003) when run on the real data, on $\mathcal{L}'$ with randomly shuffled success labels on the nodes, and on randomized networks where the node degrees are approximately preserved in expectation. As expected, for both classifiers, the mean CV accuracies for the shuffled labels are around 0.5, which corresponds to the accuracy of the naïve baseline predictor which always predicts the majority class. The mean CV accuracies obtained upon network randomization only slightly exceed 0.5 (the small improvement w.r.t. the baseline predictor can be explained by the fact that the network randomization approximately preserves the node degrees). In contrast, mean CV accuracies for the real success labels are between 0.6 and 0.7.

To contextualize the significance of the obtained mean CV accuracies for the real data, it is important to again stress that they were obtained using out-of-the-box node classifiers: We did neither tune the hyper-parameters of the employed label propagation methods, nor did we run them on a wide variety of different network models of our meta-data records. In view of this, the fact that the mean CV accuracies obtained for the real data

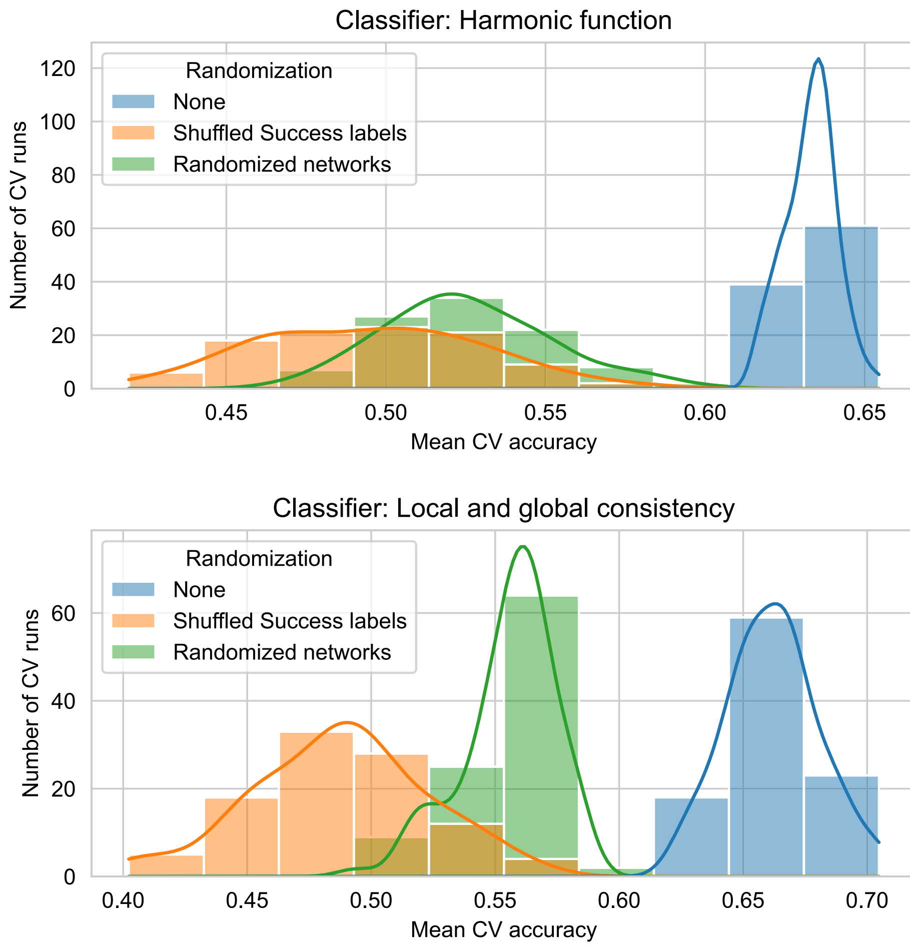

Fig. 2 Results of semi-supervised analysis on undirected line graph $\mathcal{L}'$ using the harmonic function classifier by Zhu et al. (2003) and local and global consistency classifier by Zhou et al. (2003). For comparison, the same classifiers were run on $\mathcal{L}'$ with shuffled success labels and on randomized networks where the node degrees of $\mathcal{L}'$ are approximately preserved in expectation

clearly exceed those obtained for the two random background models clearly indicates that success status is indeed encoded in the topology of our content-agnostic network models.

Node centralities characterize successful recommendation letters

Figure 3 shows the results of the differential analyses of in-centrality measures (see Table 3 for P-values underlying the significance notations). Recall that, w.r.t. these centrality measures, nodes receive large centralities if they are highly connected as recipients of recommendation letters. We make two main observations: Firstly, for all four centrality measures, sender nodes of successful letters have significantly larger in-centralities than sender nodes of unsuccessful ones. In the analyzed German-Jewish recommendation letter network, recommendation letters hence had a higher likelihood of success if the recommending individuals had themselves been recipients of many other letters. In the Discussion, we speculate on historical rationales that might explain this finding.

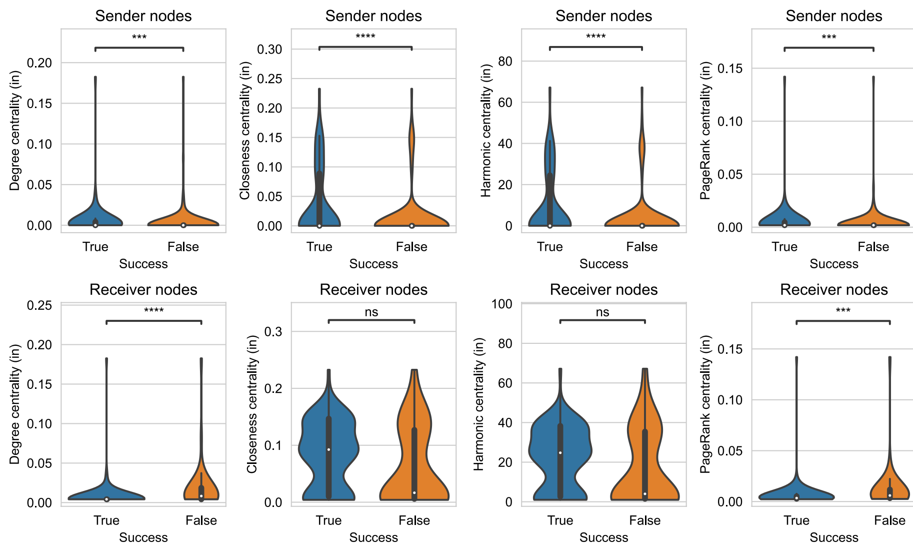

Fig. 3 Results of differential analyses of in-centralities. For all tested centrality measures, sender nodes of successful recommendation letters have significantly larger in-centralities than sender nodes of unsuccessful letters (two-sided MWW test). In contrast to that, receivers of successful recommendations have significantly smaller degree and PageRank in-centralities than receivers of unsuccessful letters (no significant differences for the closely related closeness in-centralities and harmonic in-centralities). See Table 3 for *P*-values underlying the significance annotations. Note that closeness centrality and harmonic centrality are closely related, which explains the very similar results obtained for these two centrality measures

Table 3 *P*-values of two-sided MWW tests underlying the significance annotations in Figs. 3 and 4

	Degree centrality	Closeness centrality	Harmonic centrality	PageRank centrality
MWW *P*-values obtained for in-centralities				
Senders	1.544×10^{-4}	8.863×10^{-5}	8.793×10^{-5}	1.496×10^{-4}
Receivers	3.779×10^{-5}	1.269×10^{-1}	1.641×10^{-1}	2.078×10^{-4}
MWW *P*-values obtained for out-centralities				
Senders	3.476×10^{-3}	1.160×10^{-1}	5.273×10^{-2}	1.183×10^{-8}
Receivers	2.341×10^{-1}	1.727×10^{-1}	1.818×10^{-1}	3.209×10^{-1}

The second observation is that, for two out of four centrality measures, receiver nodes of successful letters have significantly smaller in-centralities than receiver nodes of unsuccessful letters. That is, persons targeted by successful letters were on average less sought after than persons targeted by unsuccessful ones. An intuitive explanation for this result is that persons who received a lot of letters had less resources to secure positions at the targeted host institutions for the individual recommendees. For a more detailed historical contextualization, we again refer to the Discussion.

The results for the differential analyses out-centrality measures are visualized in Fig. 4. We observe that successful recommendation letters are characterized by significantly larger degree and PageRank out-centralities of sender nodes than unsuccessful letters. That is, the likelihood of a letter being successful was larger for letters that were sent by individuals who sent a lot letters (and were themselves targeted by other active senders). *Prima facie*, there are several possible explanations for this result. For instance, over time, active recommenders might have become more proficient at writing the letters,

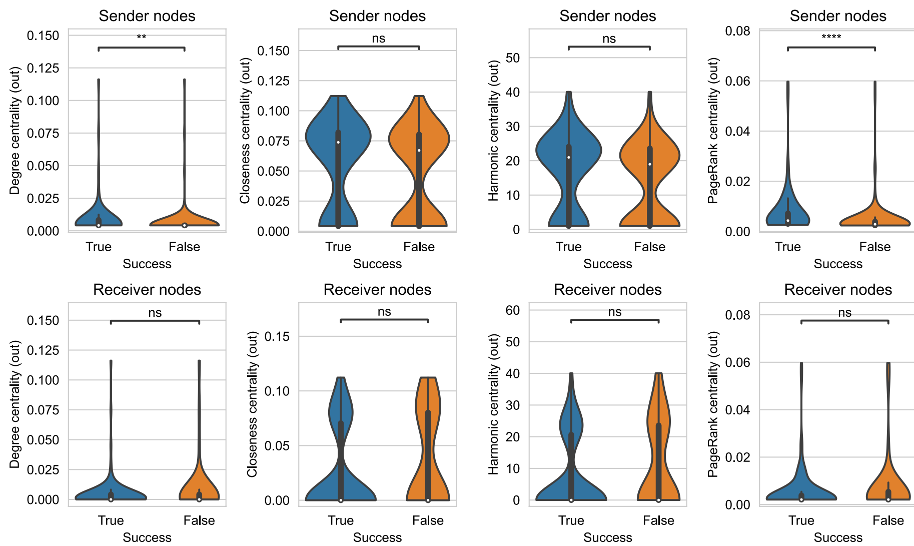

Fig. 4 Results of differential analyses of out-centralities. Sender nodes of successful recommendation letters have significantly larger degree and PageRank out-centralities than sender nodes of unsuccessful letters (two-sided MWW test). For receiver nodes, there are no significant differences. See Table 3 for *P*-values underlying the significance annotations. The very similar results for the closeness centrality and the harmonic centrality can again be explained by the fact that these two centrality measures are closely related

or recommendees might have actively addressed recommenders they knew to have authored successful recommendation letters before. A more detailed contextualization can again be found in the Discussion. For receiver nodes, we observe no significant differences.

Summary of the results

In sum, our results clearly show that the topologies of our content-agnostic German-Jewish recommendation letter networks indeed contain information about whether or not letters were successful at securing a position for the recommendee. In particular, we observe that it was important that the recommending individuals were well connected within the overall network. In contrast to that, letters tended to be more successful if they addressed individuals with lower in-connectivity—potentially, because those persons could invest more resources for the individual recommendees.

Discussion

Our findings indicate that obtaining a position at the young HU in the early 20th century was highly dependent on the academic standing and connectedness of the recommending person. Since also requests for recommendations were included to construct our network models, our first finding—namely, that recommendations had a higher success rate if the recommender also received a lot of letters—highlights the importance of connections between the HU and the recommender which had been established prior to the opening of an academic position. When requesting recommendation letters for potential candidates, the hiring committees at the HU would turn to people they knew to narrow the field of applicants. If a person approached by such a request would know a potential candidate, this candidate was more likely to be successful. This was especially the case for well-established academics like the Utrecht-based physicist Leonard Salomon Ornstein. Ornstein himself was a desired candidate by the HU, but instead of moving to Jerusalem himself, he successfully recommended a few younger colleagues (Gerlach 2022).

This interpretation is further supported by our finding that successful recommendations tended to be sent by individuals who sent many recommendation letters and were themselves targeted by active senders. The HU requested individuals to provide recommendation letters multiple times if they had provided successful recommendations in the past. Moreover, highly active and successful recommenders did not only develop a certain proficiency in writing recommendation letters, but were also experienced in pre-selecting and recommending the most promising candidates for the HU. For instance, the biologist Otto Warburg met with potential candidates in Berlin and recommended the strongest candidates based on his impressions during the meetings (Gerlach 2022).

Our third finding is that letters sent to highly sough-after individuals had lower success rates. The young HU in Jerusalem only had few available positions and was targeted by a growing number of scholars who tried to leave Germany and Western Europe due to an increase of antisemitism. The data shows how some actors received a high number of recommendations with a low hiring rate. It is very likely that these letters were addressed to persons who were well known by outsiders but without direct influence on the individual hiring processes at HU. A prime example of such a person is Judah Magnes. Between 1925 and 1935, Magnes served as chancellor and from 1935 until his death in 1948 as president of the HU. Hence, he was very visible for candidates who sought to secure a position at the HU. However, he only had limited influence on the hiring processes in the individual departments (Gerlach 2022). In our results (Fig. 5), these aspects are reflected by the fact that all four in-centrality measures are maximal for Magnes. At the same time, Magnes' success rate as a receiver of recommendation letters was significantly lower than average.

In conclusion, our study showcases how quantitative analysis of professional recommendation networks via standard network analysis methods can yield novel historical insights. Our analyses clearly show that the outcomes of applications at the HU were not solely dependent on the candidates' merits, accomplishments, and motivations, but also strongly depended on the social networks supporting their potential

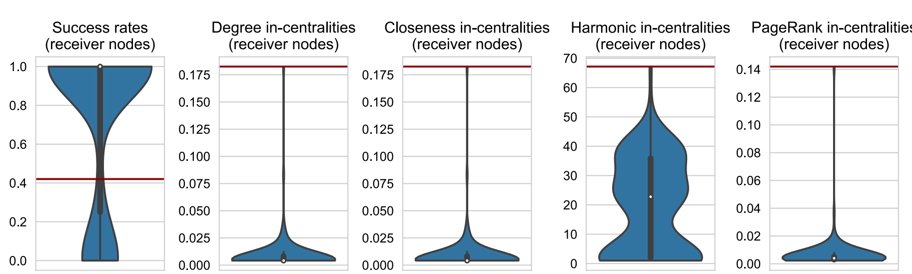

Fig. 5 Distributions of success rates, in-degree centralities $D(u)$, closeness in-centralities $CC(u)$, harmonic in-centralities $H(u)$, and PageRank in-centralities $PR(u)$, for all receivers $u \in R$. Red horizontal lines show the results for chancellor Judah Magnes. While Magnes was an extremely sought-after contact person at the HU, his success rate of 0.42 was clearly below average (mean success rate for receiver nodes: 0.73; median success rate for receiver nodes: 1.0)

employments. A successful recommendation needed to be sent by the right person and to the right person. In particular, our results indicate that an established connection between the hiring institution and the recommender was helpful and that a targeted selection of the addressee at the HU increased the likelihood of success.

In future work, it would be interesting to extend our analyses with a temporal dimension. Here, an interesting question is whether historical events or developments such as the dramatically increasing persecution of German-Jewish academics in the 1930s are reflected in structural changes in the recommendation letter networks. One approach to answer this question could be to generate temporally evolving snapshots of the recommendation letter networks via a sliding window approach and to then track changes of network or node properties over time. In this context, it might also be interesting to supplement the computation of node centralities with metrics such as the clustering coefficient (Fagiolo 2007) or the local clustering coefficient rank (Velichety and Ram 2021) which quantify structural connectivity within sub-communities.

Acknowledgements
Not applicable.

Author contributions
LG and DBB conceived and designed this study and wrote the article. LG manually generated the meta-data records and provided the historical contextualizations and interpretations. DBB implemented CorrNet and carried out the computational analyses. All authors read and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. LG was supported by travel grants and short-term fellowships from the German Historical Institute in Washington, D.C, from the Research School of the Ruhr-Universität Bochum, and from the German Research Foundation (SFB 1150, "Kulturen des Entscheidens"). We acknowledge financial support by Deutsche Forschungsgemeinschaft and Friedrich-Alexander-Universität Erlangen-Nürnberg within the funding programme "Open Access Publication Funding".

Availability of data and materials
All source code and data underlying this study are freely available on GitHub under the terms of the GNU General Public License, Version 3: https://github.com/bionetslab/corrnet. The GitHub repository also contains a CSV file with all 903 manually curated meta-data records. The original letters can be found at the Central Zionist Archive (Jerusalem), at the National Library of Israel (Jerusalem), at the he Yad Weizmann Archive (Rehovot), Baker Library Special Collection, Harvard Business School (Boston), at the Archives of the Bank Rothschild Frères (Paris), at the Sächsisches Staatsarchiv, Hauptstaatsarchiv Dresden (Dresden), at the Nietzsche Haus (Sils), at the Siemens Historical Institute (Berlin), at the Landesarchiv Baden-Württemberg, Hauptstaatsarchiv Stuttgart (Stuttgart), and at the United States Holocaust Memorial Museums Archives (Washington, D.C.).

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 22 December 2022 Accepted: 1 May 2023
Published online: 15 May 2023

References

Althage M, Dröge M, Hiltmann T, Schneider P (2022) Python fuer Historiker-innen: Publication of version 1.0. https://doi.org/10.5281/ZENODO.6868043. https://digital-history-berlin.github.io/Python-fuer-Historiker-innen/home.html

Austin AW (2012) Quaker brotherhood: interracial activism and the American Friends Service Committee, 1917–1950. University of Illinois Press, Urbana

Bavelas A (1950) Communication patterns in task-oriented groups. J Acoust Soc Am 22(6):725–730. https://doi.org/10.1121/1.1906679

Bentwich N (1953) The Rescue and Achievement of Refugee Scholars. The Story of Displaced Scholars and Scientists 1933–1952. Springer, Dordrecht

Bentwich N (1961) The Hebrew University of Jerusalem 1918–1960. Weidenfeld & Nicolson, London

Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Comput Netw ISDN Syst 30(1):107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Chung F, Lu L (2002) Connected components in random graphs with given expected degree sequences. Ann Comb 6(2):125–145. https://doi.org/10.1007/pl00012580
Dauben JW, Robinson A (1995) Abraham Robinson: the creation of nonstandard analysis: a personal and mathematical odyssey. Princeton Legacy Library, Princeton University Press, Princeton
Erel S (1983) Neue Wurzeln: 50 Jahre Immigration deutschsprachiger Juden in Israel. Bleicher, Gerlingen
Fagiolo G (2007) Clustering in complex directed networks. Phys Rev E 76(2):026,107. https://doi.org/10.1103/PhysRevE.76.026107
Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40(1):35–41. https://doi.org/10.2307/3033543
Gerlach L (2022) Empfehlungskultur: Eine historische Untersuchung zur Transformation schriftlicher Empfehlungen im Spiegel deutsch-jüdischer Geschichten zwischen 1848 und 1945. Ph.D. thesis, Ruhr-Universität Bochum, Bochum
Hagberg AA, Schult DA, Swart PJ (2008) Exploring network structure, dynamics, and function using NetworkX. In: Varoquaux G, Vaught T, Millman J (eds) Proceedings of the 7th Python in science conference, Pasadena, CA USA, pp 11–15
Lässig S (1997) Kultur und Kommerz: Das Beispiel der Bankiersfamilie Arnhold. Dresdner Hefte Beiträge zur Kulturgeschichte 15(49):39–46
Marchiori M, Latora V (2000) Harmony in the small-world. Physica A 285(3):539–546. https://doi.org/10.1016/S0378-4371(00)00311-3
Miller JC, Hagberg AA (2011) Efficient generation of networks with given expected degrees. In: Frieze AM, Horn P, Pralat P (eds) WAW 2011, Springer, Berlin, LNCS, vol 6732, pp 115–126. https://doi.org/10.1007/978-3-642-21286-4_10
Stern F (2000) Gold und Eisen: Bismarck und sein Bankier Bleichröder, neuausg., 20. - 22. tsd edn. Rowohlt-Taschenbuch-Verl., Reinbek bei Hamburg
Velichety S, Ram S (2021) Finding a needle in the haystack: recommending online communities on social media platforms using network and design science. J Assoc Inf Sci Technol 22(5):1285–1310. https://doi.org/10.17705/1jais.00694
Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat I, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, SciPy 10 Contributors (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17(3):261–272. https://doi.org/10.1038/s41592-019-0686-2
Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511815478
Zhou D, Bousquet O, Lal T, Weston J, Schölkopf B (2003) Learning with local and global consistency. In: Thrun S, Saul L, Schölkopf B (eds) Advances in neural information processing systems, vol 16. MIT Press, Cambridge
Zhu X, Ghahramani Z, Lafferty J (2003) Semi-supervised learning using gaussian fields and harmonic functions. In: Proceedings of the twentieth international conference on international conference on machine learning. AAAI Press, pp 912–919. https://www.aaai.org/Papers/ICML/2003/ICML03-118.pdf

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com