The Structure of Social Influence in Recommender Networks

Pantelis P. Analytis* pantelis@sam.sdu.dk
University of Southern Denmark

Philipp Lorenz-Spreen lorenz-spreen@mpib-berlin.mpg.de
Max Planck Institute for Human Development

Daniel Barkoczi daba@sam.sdu.dk
University of Southern Denmark

Stefan M. Herzog herzog@mpib-berlin.mpg.de
Max Planck Institute for Human Development

ABSTRACT
People’s ability to influence others’ opinion on matters of taste varies greatly—both offline and in recommender systems. What are the mechanisms underlying these striking differences? Using the weighted $k$-nearest neighbors algorithm ($k$-$nn$) to represent an array of social influence, we show—leveraging methods from network science—how the $k$-$nn$ algorithm gives rise to networks of social influence in six real-world domains of taste. We show three novel results that apply both to offline advice taking and online recommender settings. First, influential individuals have mainstream tastes and high dispersion in their taste similarity with others. Second, the fewer people an individual or algorithm consults (i.e., the lower $k$ is) or the larger the weight placed on opinions of more similar others, the smaller the group of people with substantial influence. Third, the influence networks emerging from deploying the $k$-$nn$ algorithm are hierarchically organized. Our results shed new light on classic empirical findings in communication and network science and can help improve the understanding of social influence offline and online.

KEYWORDS
social influence, influencers, social networks, collaborative filtering

ACM Reference Format:
Pantelis P. Analytis, Daniel Barkoczi, Philipp Lorenz-Spreen, and Stefan M. Herzog. 2020. The Structure of Social Influence in Recommender Networks. In Proceedings of The Web Conference 2020 (WWW ’20), April 20–24, 2020, Taipei, Taiwan. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3366423.3380020

1 INTRODUCTION
We all have opinions on matters of taste. Whether it is a new song, the design of a building, or the performance of an actor, people are eager to express their opinions offline and online. However, the opinions of some are sought out and appreciated more than the opinions of others. Consider renowned film critics such as Roger Ebert or wine critics like Robert Parker: their opinions are recognized as an indicator of quality by most other critics and the general public alike—and can thus affect the price or financial success of a product [1, 5]. Relative to such highly influential individuals, most people exert little social influence over others.

Sociologists and communication scientists have been interested in the study of influential individuals since the mid-20th century, and understandably so. By accurately identifying individuals with influence, policy makers can sway public opinion on critical matters such as public health and the diffusion of socially beneficial innovations. Early studies [27, 33, 55] surveyed large numbers of people, typically residents of representative mid-sized cities in the United States, and asked them whom they would consult for advice in various domains (e.g., public health, fashion, politics). This early work revealed (i) that within each domain people seek advice from a small group of other individuals (typically around 5), (ii) that some key individuals, commonly referred to as opinion leaders, are consistently sought out by others for advice, and therefore exert a much larger influence than others, and (iii) that opinion leaders are domain-specific. Although this work revealed that people rely on just a few individuals to inform their opinion, there was no way to evaluate the quality of the decisions that people made.

With the advent of computational methods, network theory, and the Internet, the research focus shifted to describing networks of social influence and developing methods for leveraging the clout of influential individuals in them [3, 28, 35, 54]. Social networks could be directly reconstructed by observing friendships or follower counts on online websites. Seminal methods for ranking search results, such as PageRank, use a network’s structure to assign value to different sources of information or individuals (e.g., webpages or blogs, see [41]). PageRank’s general approach has been used by social scientists to assign status to different people or sources of information in the offline world. Here, social influence is a consequence of the network’s structure, where well-connected (or well-positioned) individuals are most influential [13, 26].

Coming to grips with the structure of social influence is crucial for the recommender systems and computational social science communities. Classic collaborative filtering algorithms, such as the weighted $k$-nearest neighbors algorithm ($k$-$nn$), essentially distribute social influence among the individuals in the system’s knowledge base [11]. For each target individual, $k$-$nn$ pays attention to only a relatively small number of similar others (typically between 10 and 50, see [22, 23])—implying a particular network of social influence [29, 32]. Critically, $k$-$nn$ can also represent a broad array of decision strategies that have been studied by social and
behavioral scientists in offline settings (see Table 1). As in the communities studied by sociologists and communication scientists since the 1950s, the opinions of a few, influential individuals might be consulted more often by recommender systems. Going beyond previous research, we can now uncover the statistical properties of the opinions of the individuals whose advice is sought, and investigate the performance of different social learning strategies.

Previous research on social influence in recommender systems has focused on two main topics. First, motivated by the threat of malicious attacks on recommender systems (i.e., “shilling attacks”), researchers have developed techniques to identify and avert attackers who want to exploit the system for their own benefit [31, 45, 48]. Bot attacks in which each item is rated by its average score (with some random error added to it), are particularly effective in influencing collaborative filtering recommenders [31]. Second, researchers have leveraged social influence to design more effective collaborative filtering algorithms or run more cost-efficient marketing campaigns [11, 16, 47]. By studying the structure of social influence, we hope to derive insights into how recommendation algorithms can be further improved and made resilient against attacks.

Several questions pertaining to both offline and online opinion spaces remain unaddressed: First, is it possible to identify characteristics (e.g., statistical properties) that reliably predict whether somebody is influential or has the potential to become influential within a domain? Second, how do the recommender algorithms or social learning strategies used determine the type of social influence (e.g., varying k in k-nn or the number of people asked for advice offline)? Third, what is the structure of the networks produced by k-nn and the corresponding social learning strategies? In this paper, we investigate these three questions in a diverse set of large- and small-scale datasets.

Table 1: Correspondence between the collaborative filtering algorithm parameterizations we consider (see Equations 1, 2, and 3) and the social learning and information aggregation strategies broadly studied in the social and behavioral sciences [2].

| Social learning (taste) | Social learning (objective) | Algorithm parameters | Cognitive strategies |
|-------------------------|----------------------------|----------------------|---------------------|
| Doppelgänger [2, 56]    | Follow the expert, richest [6, 30] | k = 1 and ρ = any | Take the best, single attribute [17, 25] |
| Clique [39, 56]         | Select crowd [20, 36]           | k = n and ρ = 0     | -                   |
| Weighted clique         | Weighted select crowd [51]      | k = n and ρ > 0     | -                   |
| Weighted crowd [2]      | Weighted crowd [8]              | k = N − 1 and ρ > 0 | Weighted additive [9, 42] |
| Whole crowd [2, 56]     | Averaging [24]                 | k = N − 1 and ρ = 0 | Equal weights [9, 14] |

We used the Pearson correlation coefficient as a measure of similarity (w) between two individuals i and j [23], defined as follows:

\[ w(i, j) = \frac{\sum_{m=1}^{M}(u_{im} - \bar{u}_i)(u_{jm} - \bar{u}_j)}{\sum_{m=1}^{M}(u_{im} - \bar{u}_i)^2(u_{jm} - \bar{u}_j)^2} \]  

(2)

where \( u_{im} \) is the evaluation that the target individual i gave to item m and \( u_{jm} \) is the evaluation that the jth individual gave to the same item m. \( M \) stands for the total number of items.

We use a similarity sensitivity parameter \( \rho \) that allows us to amplify or dampen the weights of different individuals [7, 40]. We directly modify the weights obtained from Eq. 2 using the following scheme:

\[ w'_j = \begin{cases} w^0_j & \text{if } w(i, j) \geq 0 \\ -|w^0_j| & \text{if } w(i, j) < 0. \end{cases} \]  

(3)

By varying \( k \) and \( \rho \), we can produce several collaborative filtering algorithms and social learning and information aggregation strategies studied in the social and behavioral sciences [2]. For instance, setting \( \rho = 0 \) and \( k = n \) gives the original formulation of k-nn, while setting \( \rho > 1 \) overweights the opinions of the individuals more similar to the target, as is common in implementations of the weighted nearest neighbors strategy in collaborative filtering [7]. In Table 1, we illustrate how different parameterizations of our model map onto different information aggregation strategies.

2.2 The datasets

We analyzed an array of datasets, including Jester, a widely studied collaborative filtering dataset on humor collected by Goldberg and colleagues [19]; datasets on visual art, architecture, and landscapes collected by Vessel and colleagues [53]; and data on the attractiveness of people’s faces collected by DeBruine and Jones [10]. The Vessel et al. and DeBruine/Jones datasets have the structure of collaborative filtering datasets and represent key domains of interest for the recommender systems community (e.g., real estate, travel, dating). Below we specify how the stimuli were selected and describe the study protocols used to elicit the ratings. In the Vessel et al. studies, participants were asked to evaluate the same images on a 7-point scale from "not aesthetically moving" to "very aesthetically moving" two or three times; we used the average evaluation across multiple ratings from the same participant.

Visual art: 24 people evaluated 109 photographs of visual art sourced from the Catalog of Art Images Online (CAMIO) and from museum collections. The collection included lesser-known artwork from a variety of periods, styles, genres, and cultural backgrounds.
Table 2: Array of datasets representing key recommendation domains. The datasets vary in terms of the amount of shared taste (defined as the average mean taste correlation with others), the degree of taste dispersion (the mean dispersion in taste similarity with others), and the number of people who evaluated the alternatives. For large datasets, we also present (in parentheses) results for a reproducible random subsample of 100 people used for plotting in Figure 1.

| Dataset                | No. people | No. items | First published | Scale  | Shared taste | Taste dispersion |
|------------------------|------------|-----------|-----------------|-------|--------------|------------------|
| Jester                 | 14,116 (100) | 100       | Goldberg et al. (2001) | -10 to 10 | 0.113 (0.110) | 0.131 (0.132)    |
| Faces                  | 2,513 (100)  | 102       | DeBruine and Jones (2017) | 1 to 7  | 0.349 (0.348) | 0.125 (0.125)    |
| Interior architecture  | 17          | 118       | Vessel et al. (2018)     | 1 to 7  | 0.158        | 0.171            |
| Exterior architecture  | 19          | 108       | Vessel et al. (2018)     | 1 to 7  | 0.172        | 0.161            |
| Landscapes             | 18          | 148       | Vessel et al. (2018)     | 1 to 7  | 0.363        | 0.138            |
| Visual art             | 24          | 109       | Vessel et al. (2018)     | 1 to 7  | 0.105        | 0.171            |

2.3 Performance of $k$-nn

For all individuals in the dataset, we calculated the performance of different versions of the weighted $k$-nearest neighbors algorithm by independently varying the value of $k$ and similarity sensitivity parameter $\rho$. To this end, we assessed the out-of-sample performance of the $k$-nn algorithm by splitting the data into two equally sized parts: training vs. test sets. We used the training set to estimate the free parameters (i.e., the correlation coefficients between each pair of individuals; see Eq. 2). We then created all possible paired comparisons between two items in the test set and used the correlations obtained from the training set to predict which items people would prefer more strongly (i.e., rate more highly) for each version of (weighted) $k$-nn (defined by its respective pair of $k$ and $\rho$). For each individual, each version of $k$-nn, and each dataset, we calculated the proportion of correct predictions across all paired comparisons in the test set. We then averaged the results across 100 simulation repetitions.

2.4 Reconstructing social influence networks

For the network analyses, we used the same procedure as described above except that we used all items in a dataset (i.e., no cross-validation procedure). We varied the value of $k$ [i.e., 2, 5, 10, and 50] and then constructed advice networks with nodes representing the different people in the dataset. While all individuals had by definition the same number of $k$ outgoing edges connecting them to other nodes, people could have a varying number of incoming edges depending on how often the recommendation algorithm sought their advice for other people. We used node strength, defined as the sum of the absolute weights assigned to each of the $k$ nearest neighbors, as a measure of social influence that naturally fits the weighted $k$-nn algorithm and weighted networks more generally [4]. This general approach can be also used with algorithms that calculate similarity between users on a dimensionally reduced space [34] or when using other observable information about individuals to calculate similarity between them [21].

3 Results

To investigate the relation between the statistical properties of people’s taste and the performance of $k$-nn, we calculated the mean taste similarity, defined as the (arithmetic) average correlation between each individual’s taste ratings and the ratings of all of their potential peers, and taste dispersion, defined as the standard deviation of those same correlations [2]. In Table 2, we also report the grand mean of those mean taste similarities (referred as shared taste) and taste dispersions for each dataset. Unless otherwise noted, we present results for $\rho = 1$. For the Jester and Faces environments, we plot the networks for a subsample of individuals from Figure 1.

3.1 Who are the most influential individuals?

The most influential individuals are also those who benefit most from weighted $k$-nn’s recommendations; the least influential individuals benefit much less (Figure 1). To quantify this relationship, we calculated Kendall’s $\tau$ between an agent’s node strength and the predictive out-of-sample performance of the $k$-nn algorithm for that individual; we found a strong relationship in all datasets (see Figure 1 for the values of $\tau$). The most influential individuals typically have mainstream tastes but also high dispersion in taste similarity with others (Figure 2). These two attributes can be used to directly predict $k$-nn’s performance for different individuals. For example, when comparing the people with the highest and lowest accuracy in the Jester and Faces datasets, differences can be as large as 30% (see also [2]). When $k = N - 1$, the finding that individuals with higher dispersion in taste similarity exert a larger influence follows...
almost directly from our definition of influence: Their opinions on average correlate more strongly with those of others—either positively or negatively. However, this relation is non-trivial for lower values of \( k \): As \( k \) decreases, people with mainstream taste but low dispersion in taste similarity do not enter the group of consulted individuals as often. They tend to be overshadowed by individuals with slightly higher correlations to the target.

3.2 Weighting and social influence distribution

The inequality in social influence stems from two distinct, but mutually compatible, aspects of how weighted \( k\text{-}nn \) works. The first is that unequal weights are assigned to different individuals (Eq. 1). The second is that—irrespective of any weights—only a few (\( k \)) individuals are considered. For \( k = N - 1 \) and \( \rho = 1 \), the only cause of influence inequality is the simple, proportional weighting (i.e., when \( \rho = 1 \)). Even in this case, where the opinion of each individual enters the calculation of recommendations, there is some inherent inequality in people’s clout due to the different extents to which their opinions correlate with those of others. To quantify this relationship, we calculated the Gini coefficients—a common measure of inequality—for each domain for \( k = N - 1 \) and \( \rho = 1 \). Even in this case, where the opinion of each individual enters the calculation of recommendations, there is some inherent inequality in people’s clout due to the different extents to which their opinions correlate with those of others. To quantify this relationship, we calculated the Gini coefficients—a common measure of inequality—for each domain for \( k = N - 1 \) and \( \rho = 1 \). The mean Gini coefficient was 0.23, with the smallest in the landscapes environment (0.13), and the largest in the Jester environment (0.34), indicating that in all domains using the correlations directly as weights produces moderate inequality of social influence. When \( \rho \) is increased, as expected, the Gini coefficient consistently increases as well, producing a mean coefficient across environments of 0.7 for \( \rho = 10 \). Overall, social influence inequality tends to be larger in taste domains with little shared taste. To see this, compare again the Jester environment, which has the second lowest shared taste (shared taste: 0.113, Gini: 0.86) with the Landscapes environment, which has the largest (shared taste: 0.363, Gini: 0.51); this result holds for all values of \( \rho \) we investigated.

3.3 Attention and social influence distribution

In many cases, people in real life and recommender systems algorithms do not pay attention to every other individual. There are good reasons for this: focusing on a subset of people, rather than taking everybody’s opinion into account, can lead to better predictive performance [23, 51]. In addition, paying attention to fewer “advisers” can reduce the effort of actively collecting and aggregating information. In other words, even if paying attention to everybody actually improved predictive performance, it may still make sense for people to pay attention to just a few individuals. Our results show that limiting attention to a few similar others can lead to substantial influence inequality. This can be seen by comparing the average Gini coefficient across environments. In the baseline case where \( k = 5 \) and \( \rho = 1 \) (see Figure 1), the mean Gini coefficient
Figure 2: Relative node strength (i.e., influence percentile) of individuals with different mean taste similarity and dispersion of taste similarity with others (for $k = 5$ and $\rho = 1$) and across environments. People with mainstream taste and high dispersion of taste similarity with others tend to be more influential than people with alternative and idiosyncratic taste.

is 0.43, which reflects substantial inequality. Influence inequality further increases as the number of individuals to which people or algorithms pay attention decreases. For example, when $k = 2$, that is, an individual or algorithm consults only two other people, the influence distributions become even more unequal with a mean Gini coefficient of 0.53. For small values of $k$, the distribution of social influence is more unequal in environments where people have high levels of shared tastes—the inverse of high $\rho$ values. This can be seen by comparing the landscapes environment with the art environment (Figure 1, Gini 0.43 vs 0.30, respectively, for $k = 5$), or the faces environment with the Jester environment (Figure 3, Gini 0.67 vs. 0.62, respectively, for $k = 5$).

3.4 Resulting network structures

To shed more light on the social networks that emerge from $k$-nn, we focused on Jester and Faces, the two large datasets in our collection, and examined the simple case of an unweighted $k$-nn algorithm ($\rho = 0$). In this case, node strength reduces to in-degree, the arguably most basic centrality measure. In our setting, in-degree represents the number of times a node (person) was sought for advice (or involved in the calculation of a recommendation). The analysis shows that in-degree varies greatly across people: For a wide range of values of $k$ there are only a few influential individuals (hubs; see Figure 3). A second metric, the local clustering coefficient—which measures the extent to which an individual’s advisers also advise each other—is inversely related to the in-degree following the power law $C(d) = d^{-\beta}$; the less influence individuals exert over others, the tighter the clusters they tend to form (see scatter plots and fit in Fig. 3). This exact relation is predicted by the hierarchical network model [43] and cannot be accounted for by other scale-free network models. This relation is stable over a wide range of values of $k$ in both datasets; it is only lost in the Jester dataset for very large values of $k$.

4 GENERAL DISCUSSION AND CONCLUSION

Roger Ebert is probably the most famous film critic in the history of film-making. His opinion was sought by scores of movie-goers and a website bearing his name is still active. But was there something special about Ebert’s opinions that made him a nationwide phenomenon in the United States and source of advice for so many people? Are there people like Ebert in recommender systems? And is it possible to identify them solely on the basis of the statistical properties of their tastes?

Our work looks at social influence in recommender systems through the lens of network theory. Hitherto, the recommender systems community has used social networks primarily as an additional source of information [18, 37, 50], and used network theory more broadly to visualize recommender systems as bipartite user–item networks (see, e.g., [57]). Here, extending early work by Lathia et al. [32], we investigated the social networks of influence produced by the weighted $k$-nearest neighbors algorithm ($k$-nn). We found that skewed social influence distributions are inherent in recommender systems and that the emerging networks are hierarchically organized. The most influential individuals (sitting on top of the hierarchies) tend to be those who benefit the most from the $k$-nn algorithm.

Previous research showed that malicious individuals can game recommender algorithms by designing bots that evaluate options
in a way that makes the evaluations appear informative to many similar others [31, 45, 48]. Our results provide an explanation for the efficiency of averaging attacks on collaborative filtering algorithms (i.e., rating each item by its average and adding some noise). Rating profiles using averaging schemes score very high, in terms of both mean taste correlation and often also dispersion of taste similarity with the crowd. If such an individual actually existed, they would be among the most influential in the settings we studied and would benefit a lot from recommendations. More broadly, our results show that it is possible to consistently identify individuals who are more likely to become influential by looking at the statistical properties of their taste.

The k-nn algorithm and its capacity to emulate different social learning strategies provides a fresh way to look at networks of social influence in the offline world. For example, our analysis points to a simple yet plausible process by which homophily [38] and opinion leaders [27] might emerge in real-world networks: People can learn more by connecting to people who are similar to them and fare better if they limit their attention to just a few similar others. The relationship between in-degree and clustering coefficient we identified is a property of many real-world networks (e.g., the WWW or protein-interaction networks; see [52]) and we found that such a network structure can also emerge when recommendation algorithms (like k-nn) create links from pairwise similarities in people’s tastes. Some networks observed in the offline world may have emerged from mechanisms akin to those we described here—further amplified or dampened by cognitive or physical limitations that people experience offline (e.g., limitations in the size of the social network they can maintain [12] or in how sensitive they are to differences in similarity [49]).

Taken together, our results show that it is possible to analyze recommender systems algorithms and their consequences at both the individual and aggregate level. The data of each individual can be seen as a unique environment with its own statistical properties, nested within a larger overarching data structure. Understanding how the data from different individuals create structure can help us unpack the workings of recommendation algorithms and lead to the development of better and more robust recommender systems.

ACKNOWLEDGMENTS

We would like to thank Sune Lehmann for helpful discussions and K. Rhett Nichols and Susannah Goss for editing this paper.
