Information cocoons in online navigation

Article

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.isci.2022.105893

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Article

Information cocoons in online navigation

Lei Hou, Xue Pan, Kecheng Liu, Zimo Yang, Jianguo Liu, Tao Zhou

liujg004@ustc.edu.cn (J.L.)
zhutou@ustc.edu (T.Z.)

Highlights

Users surf on online recommendation networks to explore information

ICs widely exist in online recommendation networks suppressing system efficiency

Application of similarity-based recommendation techniques gives rise to ICs

Introducing flexibility to recommendation prevents ICs and improves efficiency

Article

Information cocoons in online navigation

Lei Hou,[1,2,7] Xue Pan,[1,2] Kecheng Liu,[2,4] Zimo Yang,[3] Jianguo Liu,[4,5,*] and Tao Zhou[6,*]

SUMMARY

Social media and online navigation bring us enjoyable experiences in accessing information, and simultaneously create information cocoons (ICs) in which we are unconsciously trapped with limited and biased information. We provide a formal definition of IC in the scenario of online navigation. Subsequently, by analyzing real recommendation networks extracted from Science, PNAS, and Amazon websites, and testing mainstream algorithms in disparate recommender systems, we demonstrate that similarity-based recommendation techniques result in ICs, which suppress the system navigability by hundreds of times. We further propose a flexible recommendation strategy that addresses the IC-induced problem and improves retrieval accuracy in navigation, which are demonstrated by simulations on real data and online experiments on the largest video website in China. This paper quantifies the challenge of ICs in recommender systems and presents a viable solution, which offer insights into the industrial design of algorithms, future scientific studies, as well as policy making.

INTRODUCTION

The explosive development of information technologies and services, in particular the emergence of portal sites, recommender systems, search engines, and social media, has led us to a world of abundant information. We access diverse information via increasing sources, yet it is widely believed that information cocoons (ICs) are very often emerged in which we are unconsciously trapped with limited and biased information.[1] The proliferation of ICs may result in an increase in social fragmentation, polarization, and extremism, and eventually intensify segregation and threaten democracy.[1–6]

Contributing factors to ICs are various, which can be roughly classified into two categories, namely active selection and passive choice. Individuals tend to access and produce information with similar opinions but overlook different voices.[7–9] The social network formed by like-minded people is also enhancing such information segregation that individuals are more often exposed to information communicated by his/her chosen friends.[10–16] As such, each person is at risk to be positioned in virtual "cocoons" consisting of self-selected information, leading to an echo chamber effect. Although ICs induced by active selection are of one's own choice, either intentionally or unintentionally, people also struggle with ICs of passive choice. Search engines and recommender systems are nowadays widely implemented to feed information to users according to their past records. Such feed may be very homogeneous, creating filter bubbles that narrow users' navigation scopes.[17–19] For example, a news website may recommend only conservative or liberal news to a target user based on the analytical assumption of his/her political view, or recommend friends who have very similar political views. Consequently, the behaviors of active selection and passive choice may coact and reinforce ICs via friend recommendations[20,21] and news recommendations.[22–24]

Although IC-related issues are under the spotlight of investigation and heated debates,[11,25–31] quantitative studies about the existence and influence of ICs are rare, largely because of the lack of an explicit definition of IC and subsequently a benchmark for quantitative analyses. Here we provide a mathematically formal definition of IC in a common scenario of online navigation, namely the recommendation network (RN) that connects similar contents with hyperlinks (URL links) according to algorithmic evaluations[32–34] (Figure 1, Figure S1). Denoting $G(V, E)$ a directed RN where V and E are a set of nodes (objects) and a set of directed links (hyperlinks), then an IC is defined as a subset $C \in V$ such that (1) the subgraph $G[C]$ induced by C is strongly connected, and (2) there is no outgoing link from a node in C to a node outside C. A node belonging to an IC is called an IC node (note, a node at most belongs to one IC) and the number of nodes in an IC is called its size.

[1]School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

[2]Informatics Research Centre, University of Reading, Reading RG66UD, UK

[3]Beijing AiQiYi Science & Technology Co. Ltd., Beijing 100080, China

[4]Institute of Accounting and Finance, Shanghai University of Finance and Economics, Shanghai 200433, China

[5]Research Group of Computational and AI Communication at Institute for Global Communications and Integrated Media, Fudan University, Shanghai 200433, China

[6]Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, China

[7]Lead contact

*Correspondence:
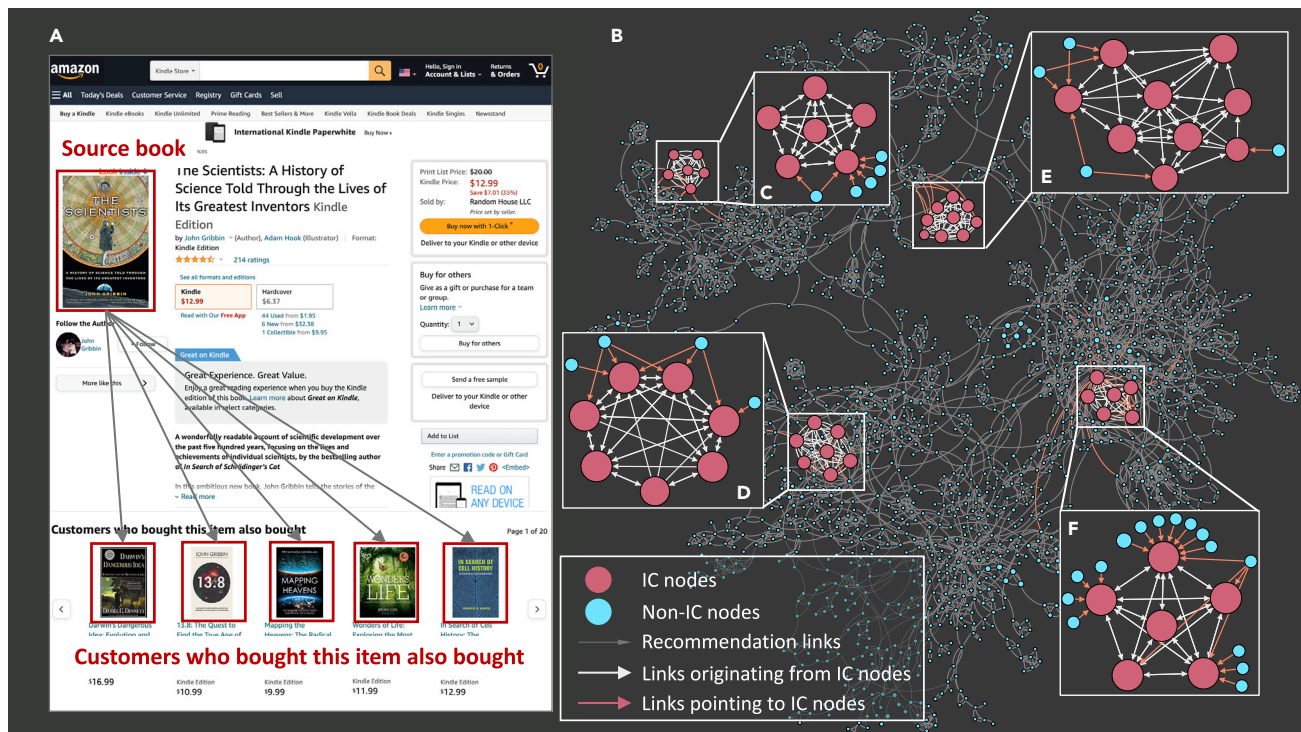liujg004@ustc.edu.cn (J.L.),
zhutou@ustc.edu (T.Z.)

https://doi.org/10.1016/j.isci.2022.105893

Figure 1. The recommendation network (RN) of *Amazon* kindle books

(A) Screenshot of a book's webpage from *Amazon*, where a list of recommended books is displayed with hyperlinks (URL links) embedded. More examples are shown in Figure S1.

(B) A sample of collected *Amazon* RN, where each node is a kindle book, and each directed link represents a hyperlink. The node sizes are proportional to the logarithm of visiting frequencies of a random walk.

(C–F) Showcases of four information cocoons and their neighboring nodes, which are empirically observed in the *Amazon* RN.

With such a definition, the aim of this paper is threefold: (1) To quantify the impact of ICs on the efficiency of online navigation systems; (2) to unfold the mechanism underlying the emergence of ICs; and (3) to provide a solution to ease IC-induced problems.

Impact of ICs on navigability

We firstly examine three empirical RNs, namely the *Science* RN of articles, *PNAS* RN of articles, and *Amazon* RN of kindle books, which are collected from the three mentioned websites (see method details for description of data collection). Figure 1 illustrates the case of *Amazon*. In the website of *Amazon*, the page of each kindle book lists several recommended books with hyperlinks embedded (Figure 1A). These hyperlinks constitute the *Amazon* RN (Figure 1B), aiming to help users explore relevant information. Table 1 presents fundamental statistics of the three empirical RNs. In such RNs, the in-degree of a node k largely describes how often the article/kindle book gets recommended by others, and thus well links with its visibility in the network for the surfing users. As shown in Figure 2A, the in-degree distributions of empirical RNs show heavy-tailed patterns. Although most nodes barely get recommended, there are hub nodes that frequently show up in others' recommendation lists, and thus have much higher chances to be visited by users. In particular, according to the definition, 96, 79, and 1181 ICs in *Science*, *PNAS*, and *Amazon* RNs have been identified, respectively (Figures 1C–1F reveal four typical ICs in *Amazon* RN, see method details for the identification of ICs). Most ICs in empirical RNs are of rather small sizes (Table S1). This is largely owing to the strict definition, because a large subgraph is unlikely to be strongly connected.

We apply randomwalks[35] to simulate users' surfing activities. Generally, the more nodes being visited within a given number of clicks, the more diverse information could be accessed. Such a quantity can be well characterized by network navigability.[36] Given an RN with N nodes, its navigability $\Omega(G)$ can be defined as the expected coverage of distinct nodes being visited during an N-steps random walk from a randomly selected starting node. Accordingly, a higher navigability suggests higher diversity of information access

Table 1. Statistics for the studied recommendation networks

RN	#Objects	#ICs	#IC nodes	IC traffic	Navigability
Empirical RNs					
Science	7,730	96	350	94.77%	0.44%
PNAS	59,479	79	415	96.98%	0.67%
Amazon	119,636	1,181	10,859	95.81%	0.07%
Derived RNs					
Steam	10,978	10.00	113.70	99.99%	0.06%
Yelp	60,785	8.90	164.90	22.85%	0.09%
Epinions	61,273	3.00	46.30	99.98%	0.05%
MovieLens	33,670	1.00	8.00	99.99%	0.03%

The symbol # stands for the number of and IC traffic means the percentage of visits on IC nodes during an N-steps random walk. The results regarding derived RNs are averaged over 20 realizations, and the standard deviations among different realizations are reported in Table S3.

from the RN. Denoting $n(t)$ the expected number of distinct nodes being visited during a t-steps random walk, for a completely random network, the growth of $n(t)$ follows the dynamics

$$\frac{d}{dt}n(t) = 1 - \frac{1}{N}n(t),$$ (Equation 1)

and, thus, we have

$$n(t) = N\left(1 - e^{-t/N}\right)$$ (Equation 2)

Hence, the corresponding navigability is

$$\Omega = \frac{n(N)}{N} = 1 - \frac{1}{e} \approx 63.21\%.$$ (Equation 3)

To validate the above prediction, we create random RNs with $N = 5 \times 10^4$ nodes by letting each node connects to $L = 5$ others randomly. As shown in Figure 2B, simulations in random RNs well follow such prediction. However, to our surprise, navigabilities of the three empirical RNs are all less than 1% (see Figure 2C), whereas IC nodes monopolize most traffic (generally $\geq$ 95%, see Table 1).

As in-degree distributions of the three RNs are heavy-tailed, it is also possible that hub nodes with large in-degrees dominate the traffic. To separate effects from ICs and hub nodes, we apply the link-crossing operations sufficiently many times to get first-order null networks.[37] In each operation, two links, say $a \to b$ and $c \to d$, are randomly selected and switched as $a \to d$ and $c \to b$. The selection ensures the avoidance of multiple links and loops. In a null network, the degree sequence keeps unchanged while ICs are absent. As shown in Figure 2C, in despite of the presence of hub nodes, $n(t)$ curves for null networks closely follow the prediction of random networks, suggesting that IC nodes rather than hub nodes result in poor navigabilities.

To further demonstrate the impact of ICs on navigability, we insert ICs to completely random RNs where each node connects to L random others. To insert an IC to the RN, we randomly select a node as the target, remove all out-going links of its L recommending nodes, and reconnect the target and its L recommending nodes to form a fully connected network. Then we obtain an IC of size $S = L + 1$. By manipulating the number of inserted ICs, the number of IC nodes, denoted as c, can be controlled accordingly. Assuming $c > 0$ IC nodes are inserted into a random RN. The random walk in this network can thus be regarded as a Bernoulli process, where at each step, the walker has a probability of $\rho = c/N$ to visit an IC node, and thereby fall into an IC. The number of steps until falling into an IC for the first time, denoted as s_0, follows a Geometric distribution as $p(s_0 = t) = \rho(1 - \rho)^t$. Consequently, the expected number of steps until falling into an IC is $\langle s_0 \rangle = 1/\rho = N/c$. Once the walker falls into an IC, only the nodes within this IC can be visited. Therefore, the number of distinct visited objects during an N-steps random walk can be calculated by summing up two parts: visited nodes before and after falling into an IC. Accordingly, we have

$$n(N) = n(t = s_0) + S - 1 = N\left(1 - e^{-\frac{s_0}{N}}\right) + S - 1.$$ (Equation 4)
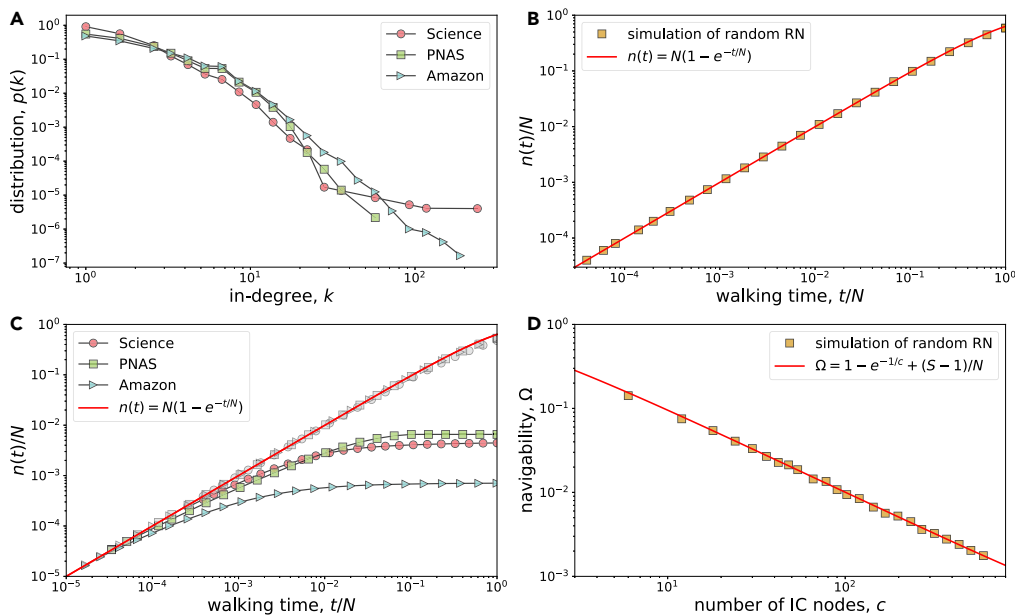
Figure 2. Structural efficiency of empirical RNs

(A) Binned in-degree distributions of empirical RNs.

(B and C) Number of distinct nodes being visited during random walks in the random RNs and empirical RNs respectively. The red curve denotes the prediction from a completely random network (Equation 2). Gray circles, squares, and triangles in (C) represent the results for the null networks of *Science*, *PNAS,* and *Amazon* RNs.

(D) Navigability of random RNs with manipulated ICs. The red curve denotes the prediction as specified by Equation 5. All the reported results regarding random RNs are averaged over 20 independent realizations with $N = 5 \times 10^4$ nodes and each node connecting to $L = 5$ random others. For each network under consideration, the result is averaged over N random walk experiments with each node being once the starting node.

Taking $s_0 = N/c$ into the above equation, the navigability of a random RN with c IC nodes is thus

$$\Omega = \frac{n(N)}{N} = 1 - e^{-\frac{1}{c}} + \frac{S-1}{N}. \qquad \text{(Equation 5)}$$

As shown in Figure 2D, the simulation in random RNs with manipulated ICs suggests the rapid decrease of navigability as predicted. Impressively, with even one IC inserted, the navigability dramatically drops to 14.19%. With more ICs being inserted, the navigability further decreases. Therefore, we conclude that the existence of ICs largely causes to the poor navigability of RNs.

Mechanism to form ICs

Though it is very likely that links in empirical RNs connect similar objects, we do not know the exact mechanism underlying empirical RNs. Therefore, we next generate RNs by implementing mainstream recommendation algorithms based on datasets of real user-object interactions. We consider four real datasets (Steam, Yelp, Epinions, and MovieLens, see method details), each of which can be described by a bipartite network $G^B(U, O, E^B)$ where $U = \{u_1, u_2, \cdots, u_M\}$ is the set of users, $O = \{o_1, o_2, \cdots, o_N\}$ is the set of objects, and E^B is the set of links between users and objects.[38,39] According to many widely applied similarity-based recommendation techniques, a recommendation network G can be generated by linking each object to its top-L most similar objects with pairwise similarity being defined based on G^B. We adopt the common neighbor index[40–42]

$$s_{\alpha\beta} = \sum_{u \in U} b_{u\alpha} b_{u\beta} + \epsilon, \qquad \text{(Equation 6)}$$

where $\epsilon \to 0$ is a tiny random number used to remove degeneracy caused by same similarity scores and $B^{(M \times N)}$ is the adjacency matrix of G^B with $b_{uo} = 1$ if user u connects with object o and $b_{uo} = 0$ otherwise.

Analogous to the empirical RNs, all four derived RNs have heavy-tailed in-degree distributions (see Figure 3A), suggesting that the similarity-based recommendation technique tends to emphasis on some
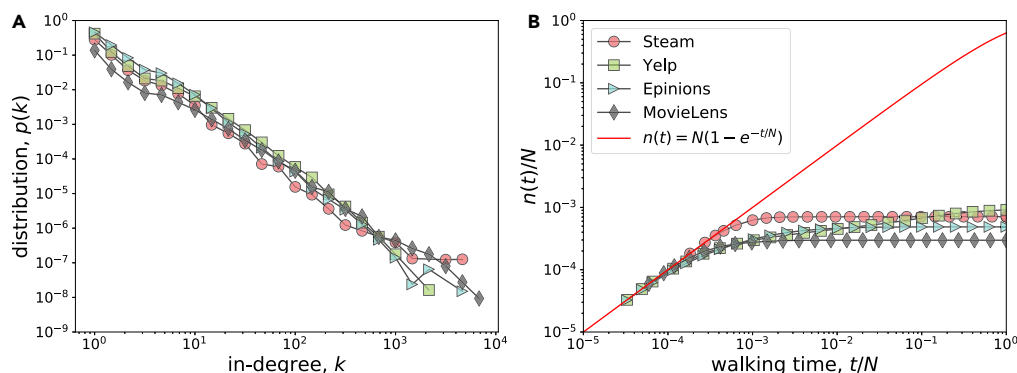
Figure 3. Structural efficiency of derived RNs
(A) Binned in-degree distributions and (B) Number of distinct nodes being visited during random walks in the derived RNs
for $L = 5$.

particular objects, making them frequently recommended. A few ICs also emerged in derived RNs. Though generally with very small sizes (Table S2), these ICs monopolize a significantly large amount of traffic (see Table 1). In particular, as shown in Figure 3B, derived RNs have much lower navigabilities in comparison with random RNs. An N-steps random walk can only find 0.06%, 0.09%, 0.05%, and 0.03% objects in Steam, Yelp, Epinions, and MovieLens RNs respectively.

The results for other well-known similarity indices are close (see Table S3 for results of Jaccard index,[41,42] Salton index,[41,42] and heat conduction index[17,43]). In a word, the similarity-based recommendation algorithms can generate ICs and thus lead to poor navigability.

Flexible recommendation

A possible cause of ICs is the similarity reciprocity (i.e., if α is among the most similar objects to β, then β is likely among the most similar objects to α) and similarity transitivity (i.e., if both α and β are among the most similar objects to γ, then α and β are likely to be very similar to each other). This subsequently leads to the formation of local clusters if we simply pick up the top-L most similar objects to construct the RN, and ICs are an extreme type of such clusters. To break ICs and thus improve navigability, we suggest a flexible recommendation strategy that selects the L recommended objects of each object from its top-λL ($\lambda > 1$) most similar objects (see Figure 4A for an illustration). As shown in Figures 4B, S6, and S7, the increasing λ quickly reduces ICs and largely improves navigability. Meanwhile, we should also consider the effect of λ on the ability to hit a user's interest. To quantify such ability, in each user-object interaction dataset, users are randomly divided into a training group and a testing group, and only the information of training users is used to construct the RN. Each testing user u then performs a random walk starting from one of u's selected objects. After t steps, the hit rate of u's interest is

$$r_u(t) \;=\; h_u(t) \,/\, (k_u - 1), \tag{Equation 7}$$

where k_u is the number of selected objects of u (i.e., the degree of u in the original user-object bipartite network) and $h_u(t)$ is the number of visited objects among the $k_u - 1$ selected objects (except for the starting object) during the t-step random walk. The overall retrieval accuracy $r(t)$ is the average of hit rates over all testing users. As shown in Figure 4C and Figure S8, the optimal value of λ subject to the largest $r(t)$ is larger than 1 unless t is very small, and there exists a huge area in the (λ, t) plane wherein the navigability and retrieval accuracy can be simultaneously improved.

We further tested whether the flexible recommendation strategy is effective in a real scenario of online navigation. The experiment was carried out in AiQiYi (NASDAQ: IQ), the largest video website in China with about 1.5×10^8 daily active users and 5×10^8 monthly active users (about 2/3 users use mobile app). To fill a recommendation position, relevant videos are selected by a series of recall algorithms from all candidates and then sorted by a ranking model. The top item that can pass the final regulation (to filter out violent, porno, and brand-conflicting videos) will be exhibited (see Figure S10 for an illustration of the structure of AiQiYi's recommender system). The item-based collaborative filtering (ICF) is a major recall algorithm, which finds out the most relevant videos according to recently clicked videos of the target user. Upon each
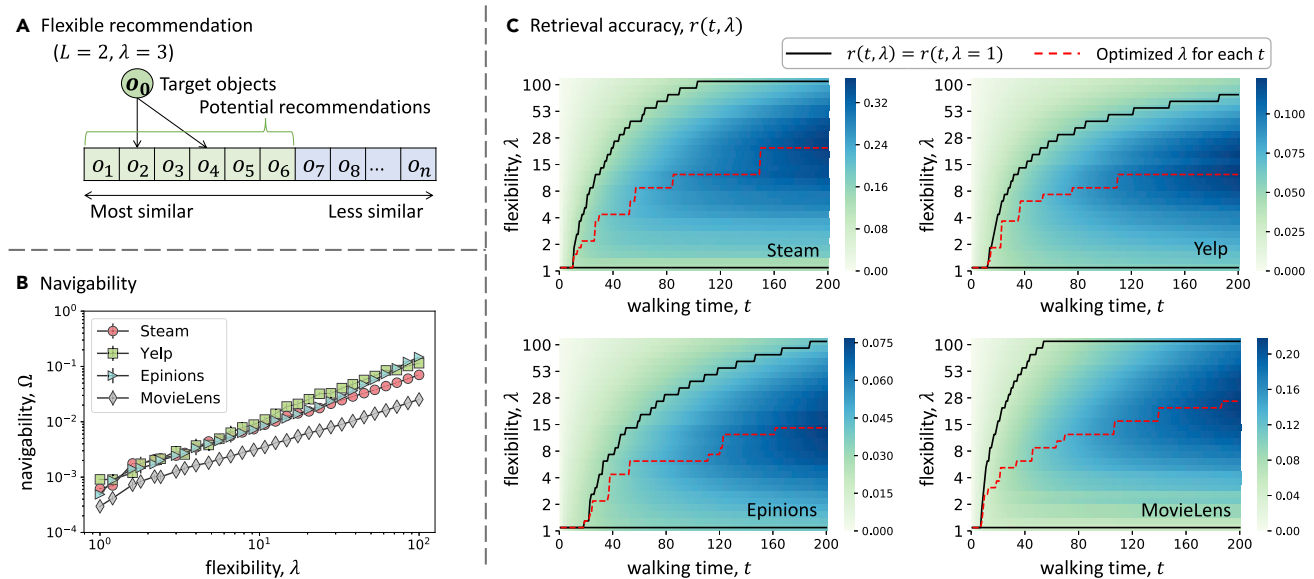
A Flexible recommendation
($L = 2, \lambda = 3$)

B Navigability

C Retrieval accuracy, $r(t, \lambda)$

Figure 4. Efficacy of the flexible recommendation strategy

(A) Illustration of the flexible recommendation strategy, where the target object randomly recommends L objects from the pool of λL most similar ones.

(B) Navigabilities of the four derived RNs with different λ.

(C) Heatmaps for retrieval accuracy in the (λ, t) plane, where the black solid curves mark the areas with improved accuracy. The red dashed lines indicate the optimal λ subject to the highest accuracy. The results reported in (B) and (C) are obtained based on the common neighbor index and a 90%–10% division of the training and testing groups. For each dataset, given the flexibility λ, the navigability is averaged over 100 realizations of RNs, and for each RN, the result is averaged over N random walk experiments with each node being once the starting node. The error bars in (B) are standard deviations of navigability across different independent experiments. The retrieval accuracy is also averaged over 100 realizations of RNs, and for each RN, 5 independent random walk experiments are carried out for each pair of a testing user and a starting object.

request, the original ICF returns the top-5 most relevant videos. In our experiment, for users in the treatment group, it randomly returns 5 videos from the top-10 most relevant ones, analogous to the flexible recommendation strategy with $\lambda = 2$. The experiment was conducted in the first two positions of the *Guess You Like* column on the landing page, which are the hottest positions attracting about 1.5×10^8 clicks from about 8×10^7 distinct users per day. To evaluate the performance, we employ two widely used metrics in industry, playing rate (PR) and playing duration (PD). The former is the ratio of playing to clicking of recommended videos, and the latter is the average playing duration (see method details). The experiment lasted one week from November 3 to November 9 in 2020 (daily results are presented in Table S4), with average PR and PD over seven days being 73.22% and 61.37 min for the treatment group (5% users, randomly selected), and 73.12% and 61.33 min for the control group (95% users). Our experiment only made a minute alteration of an elaborately designed and well trained recommender system in industry but brought about 150,000 more video plays per day (the change of PR is statistically significant, see *t*-test in method details), indicating the effectiveness of the flexible recommendation strategy.

DISCUSSION

Despite ongoing and heated debates on the harm of ICs,[11,25–31] a formal definition of IC is lacking. The primary contribution of this paper is to provide a mathematically explicit definition of IC, and to demonstrate the existence and notably negative impact on the navigability of ICs in both empirical and derived recommendation networks. The definition may appear to be too strict and thus less applicable; however, based on the essence of our research, it can be extended to characterize more generalized substructures of directed networks. For example, the extent a strongly connected subgraph $G[C]$ induced by a node set C is likely to form an IC can be measured by its escaping probability $p^e(C)$, defined as the ratio of escaping links (i.e., links from nodes in C to nodes outside C) to all links starting from nodes in C. Then, a strongly connected subgraph with a p^e no more than a preset threshold can be treated as a quasi-IC (QIC, see Figure S11). Such an escaping probability of QIC is also closely linked with the system navigability (see method details), which improves the explanation power. For example, as shown in Figure S12, the two QICs in Yelp,

respectively of escaping probabilities 0.0111 and 0.0118, dominate 71.86% of the random walk traffic. The remarkably lower IC traffic of Yelp in Table 1 can thus be well explained.

Similarity based recommendation algorithms used to be popular and are still important modules in industrial recommender systems up to date.[32,44] Present simulations on similarity-based algorithms indicate that recommender systems, by nature of their design, tend to insulate users from exposure to diverse content. Recent ethical studies[19,45,46] have noticed this issue and suggested the avoidance of filter bubbles as a high-priority task in designing or improving recommender systems, but they do not provide a viable pathway toward the target. Relevant acts and regulations emphasize algorithmic transparency (see Algorithmic Justice and Online Platform Transparency Act, as a bill in the Senate of the United States) and users' rights in shutting down recommending services and deleting part or all of personalized tags (see Management Regulations on Algorithmic Recommendations in Internet Information Services, proposed by the Cyberspace Administration of China). However, these rules cannot save us from ICs because filter bubbles are not resulted from opacity and users are usually not aware of (or even enjoying) biased information. What is worse is that such acts and regulations, if not being rightly applied, may reduce our benefits from algorithms underlying online navigation. Different from known ethical suggestions and legal rules, our results indicate that IC-related problems can be well addressed inside the algorithmic framework, using the proposed flexible strategy or other alternatives.[17,47,48] By deploying mathematical concepts, quantitative analyses, and computational tools, this paper describes, characterizes, and solves IC-related problems, which also provides a referential framework to other problems related to technical ethics.

Limitations of the study

Filter bubbles could have various causes and representations, and are not necessarily limited to object networks as studied in the present paper. For example, on the landing page of individual users (instead of when visiting a particular object), a filter bubble could be created by personalised recommendations, containing limited and biased objects, according to the system's evaluation on the target user's interest. The ICs in the present paper, on the other hand, are only an extreme form of filter bubbles in a particular scenario (surfing the recommendation network), which can be mathematically defined and evaluated. Algorithmic solutions for general forms of filter bubbles thus still need further research attention.

Owing to the diverse contexts considered (e.g., research articles, books, products), this study has only focused on the efficiency of the system to enable access to more objects. We perform a small-scale experiment with the Science dataset to illustrate that ICs are very likely to be formed by objects with similar content. For any two research articles, i and j, their content similarity can be reflected by the ratio of co-words in their titles, calculated as $R = w_{ij}^c/(w_i + w_j - w_{ij}^c)$, where w_i and w_j are the number of distinct words in i's and j's title, and w_{ij}^c is the number of co-words between the titles of i and j. Punctuations and stop words (e.g., "we", "in", "can") are not considered. Calculations based on 10^5 randomly sampled article pairs reveal that the average co-word ratio of the pairs without recommendation relation is 0.0026, whereas that of the recommended article pairs is 0.0194. Such a result suggests that each object tends to recommend others with very similar content. As a consequence, the objects in the same IC tend have also similar content. Accordingly, the content homogeneity of the accessed objects during surfing in recommendation networks shall be explored in future research, which could potentially enrich the understanding of the impact of recommendation algorithms on not only the limited but also the biased information access.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Collection of empirical recommendation networks
  - Strategy for identification of information cocoons

○ Datasets of user-object interactions
○ Setting and results of the online experiment
○ Quasi information cocoons

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105893.

AUTHOR CONTRIBUTIONS

L.H., K.L., J.L., and T.Z. conceived and designed the study. L.H. and X.P. collected and analyzed the data. Z.Y. performed the online experiment. L.H. and T.Z. drafted the manuscript. L.H., K.L., J.L., and T.Z. revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Sunstein, C.R. (2006). Infotopia: How Many Minds Produce Knowledge (Oxford University Press).

2. Stroud, N.J. (2010). Polarization and partisan selective exposure. J. Commun. 60, 556–576.

3. Sunstein, C.R. (2018). Is social media good or bad for democracy. Int. J. Hum. Rights 27, 83–89.

4. Shi, F., Shi, Y., Dokshin, F.A., Evans, J.A., and Macy, M.W. (2017). Millions of online book co-purchases reveal partisan differences in the consumption of science. Nat. Hum. Behav. 1, 0079.

5. Sülflow, M., Schäfer, S., and Winter, S. (2019). Selective attention in the news feed: an eye-tracking study on the perception and selection of political news posts on Facebook. New Media Soc. 21, 168–190.

6. Romenskyy, M., Spaiser, V., Ihle, T., and Lobaskin, V. (2018). Polarized Ukraine 2014: opinion and territorial split demonstrated with the bounded confidence XY model, parametrized by Twitter data. R. Soc. Open Sci. 5, 171935.

7. Cowan, S.K., and Baldassarri, D. (2018). "It could turn ugly": selective disclosure of attitudes in political discussion networks. Soc. Networks 52, 1–17.

8. Wihbey, J., Joseph, K., and Lazer, D. (2019). The social silos of journalism? Twitter, news media and partisan segregation. New Media Soc. 21, 815–835.

9. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., and Starnini, M. (2021). The echo chamber effect on social media. Proc. Natl. Acad. Sci. USA 118. e2023301118.

10. Hu, J., Zhang, Q.M., and Zhou, T. (2019). Segregation in religion networks. EPJ Data Sci. 8, 6.

11. Bakshy, E., Messing, S., and Adamic, L.A. (2015). Exposure to ideologically diverse news and opinion on Facebook. Science 348, 1130–1132.

12. Mosleh, M., Martel, C., Eckles, D., and Rand, D.G. (2021). Shared partisanship dramatically increases social tie formation in a Twitter field experiment. Proc. Natl. Acad. Sci. USA 118. e2022761118.

13. Chen, W., Pacheco, D., Yang, K.C., and Menczer, F. (2021). Neutral bots probe political bias on social media. Nat. Commun. 12, 5580.

14. Vasconcelos, V.V., Levin, S.A., and Pinheiro, F.L. (2019). Consensus and polarization in competing complex contagion processes. J. R. Soc. Interface 16, 20190196.

15. Tokita, C.K., and Tarnita, C.E. (2020). Social influence and interaction bias can drive emergent behavioural specialization and modular social networks across systems. J. R. Soc. Interface 17, 20190564.

16. Ou, Y., Guo, Q., and Liu, J. (2022). Identifying spreading influence nodes for social networks. Front. Eng. Manag. 9, 520–549.

17. Zhou, T., Kuscsik, Z., Liu, J.G., Medo, M., Wakeling, J.R., and Zhang, Y.C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. Proc. Natl. Acad. Sci. USA 107, 4511–4515.

18. Pariser, E. (2011). The Filter Bubble: What the Internet Is Hiding from You (Penguin).

19. Helberger, N., Karppinen, K., and D'acunto, L. (2018). Exposure diversity as a design principle for recommender systems. Inf. Commun. Soc. 21, 191–207.

20. Aiello, L.M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., and Menczer, F. (2012). Friendship prediction and homophily in social media. ACM Trans. Web 6, 1–33.

21. Huszár, F., Ktena, S.I., O'Brien, C., Belli, L., Schlaikjer, A., and Hardt, M. (2022). Algorithmic amplification of politics on Twitter. Proc. Natl. Acad. Sci. USA 119. e2025334119.

22. Beam, M.A. (2014). Automating the news: how personalized news recommender system design choices impact news reception. Commun. Res. *41*, 1019–1041.

23. Santos, F.P., Lelkes, Y., and Levin, S.A. (2021). Link recommendation algorithms and dynamics of polarization in online social networks. Proc. Natl. Acad. Sci. USA *118*. e2102141118.

24. Ohme, J. (2021). Algorithmic social media use and its relationship to attitude reinforcement and issue-specific political participation–The case of the 2015 European Immigration movements. J. Inf. Technol. Politics *18*, 36–54.

25. Yang, T., Majó-Vázquez, S., Nielsen, R.K., and González-Bailón, S. (2020). Exposure to news grows less fragmented with an increase in mobile access. Proc. Natl. Acad. Sci. USA *117*, 28678–28683.

26. ZuiderveenBorgesius, F.J., Trilling, D., Möller, J., Bodó, B., De Vreese, C.H., and Helberger, N. (2016). Should we worry about filter bubbles? Internet Policy Review *5*, 1–14.

27. Guess, A., Nyhan, B., Lyons, B., and Reifler, J. (2018). Avoiding the Echo Chamber about Echo Chambers*2* (Knight Foundation), pp. 1–25.

28. Bruns, A. (2019). Are Filter Bubbles Real? (John Wiley & Sons).

29. Eady, G., Nagler, J., Guess, A., Zilinsky, J., and Tucker, J.A. (2019). How Many People Live in Political Bubbles on Social Media? Evidence from Linked Survey and Twitter Data, *9* (Sage Open), 2158244019832705.

30. Powers, E. (2017). My news feed is filtered? Awareness of news personalization among college students. Digit.Journal. *5*, 1315–1335.

31. Puschmann, C. (2019). Beyond the bubble: assessing the diversity of political search results. Digit.Journal. *7*, 824–843.

32. Lü, L., Medo, M., Yeung, C.H., Zhang, Y.C., Zhang, Z.K., and Zhou, T. (2012). Recommender systems. Phys. Rep. *519*, 1–49.

33. Oestreicher-Singer, G., and Sundararajan, A. (2012). Recommendation networks and the long tail of electronic commerce. MIS Q. *36*, 65–83.

34. Kumar, A., and Hosanagar, K. (2019). Measuring the value of recommendation links on product demand. Inf. Syst. Res. *30*, 819–838.

35. Masuda, N., Porter, M.A., and Lambiotte, R. (2017). Random walks and diffusion on networks. Phys. Rep. *716–717*, 1–58.

36. De Domenico, M., Solé-Ribalta, A., Gómez, S., and Arenas, A. (2014). Navigability of interconnected networks under random failures. Proc. Natl. Acad. Sci. USA *111*, 8351–8356.

37. Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. Science *296*, 910–913.

38. Zhou, T., Ren, J., Medo, M., and Zhang, Y.C. (2007). Bipartite network projection and personal recommendation. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *76*, 046115.

39. Shang, M.S., Lü, L., Zhang, Y.C., and Zhou, T. (2010). Empirical analysis of web-based user-object bipartite networks. Europhys.Lett. *90*, 48006.

40. Liben-Nowell, D., and Kleinberg, J. (2007). The link prediction problem for social networks. J. Am. Soc. Inf. Sci. Tec. *58*, 1019–1031.

41. Lü, L., and Zhou, T. (2011). Link prediction in complex networks: a survey. Phys. A Stat. Mech. Appl. *390*, 1150–1170.

42. Liu, J.G., Hou, L., Pan, X., Guo, Q., and Zhou, T. (2016). Stability of similarity measurements for bipartite networks. Sci. Rep. *6*, 18653.

43. Zhang, Y.C., Blattner, M., and Yu, Y.K. (2007). Heat conduction process on community networks as a recommendation model. Phys. Rev. Lett. *99*, 154301.

44. Smith, B., and Linden, G. (2017). Two decades of recommender systems at amazon.com. IEEE Internet Comput. *21*, 12–18.

45. Milano, S., Taddeo, M., and Floridi, L. (2020). Recommender systems and their ethical challenges. AI Soc. *35*, 957–967.

46. Polonioli, A. (2021). The ethics of scientific recommender systems. Scientometrics *126*, 1841–1848.

47. Harambam, J., Helberger, N., and van Hoboken, J. (2018). Democratizing algorithmic news recommenders: how to materialize voice in a technologically saturated media ecosystems. Philos. Trans. A Math. Phys. Eng. Sci. *376*, 20180088.

48. Abdollahpouri, H., Burke, R., and Mobasher, B. (2020). Managing popularity bias in recommender systems with personalized re-ranking. In Proceedings of the 32th International Flairs Conference, pp. 413–418.

49. Pathak, A., Gupta, K., and McAuley, J. (2017). Generating and personalizing bundle recommendations on steam. In Proceedings of the 40th International ACMSIGIR Conference on Research and Development in Information Retrieval, pp. 1073–1076.

50. Massa, P., and Avesani, P. (2007). Trust-aware recommender systems. In Proceedings of the 2007 ACM conference on recommender systems, pp. 17–24.

51. Harper, F.M., and Konstan, J.A. (2016). The movielens data sets: history and context. ACM Trans. Interact. Intell. Syst. *5*, 1–19.

52. Fiorini, P.M., and Lipsky, L.R. (2012). Search marketing traffic and performance models. Comput. Stand. Interfac. *34*, 517–526.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Science, PNAS, and Amazon recommendation network	This paper	https://github.com/LLLLHou/Information-Cocoons-in-online-Navigation
Steam user-object interaction data	Pathak et al. (2017)[49]	https://github.com/LLLLHou/Information-Cocoons-in-online-Navigation
Yelp user-object interaction data	Yelp dataset challenge	https://github.com/LLLLHou/Information-Cocoons-in-online-Navigation
Epinions user-object interaction data	Massa and Avesani (2007)[50]	https://github.com/LLLLHou/Information-Cocoons-in-online-Navigation
MovieLens user-object interaction data	Harper and Konstan (2016)[51]	https://github.com/LLLLHou/Information-Cocoons-in-online-Navigation
Analytical codes in C	This paper	https://github.com/LLLLHou/Information-Cocoons-in-online-Navigation
Software and algorithms		
Python version 3.7.6	Python Software Foundation	https://www.python.org
XCode version 14.1	Apple	https://developer.apple.com/xcode/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Lei Hou (l.hou@nuist.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data of empirical recommendation networks and user-object bipartite networks have been deposited at GitHub and are publicly available as of the date of publication. The link is provided in the key resources table. The experimental data of AiQiYi is not openly accessible due to privacy concerns.

- All original code has been deposited at GitHub and is publicly available as of the date of publication. The link is provided in the key resources table.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Collection of empirical recommendation networks

In many online content systems, each object (e.g., product, article, movie, etc.) is presented on a designated webpage. In such a webpage, in addition to the primary description, a recommendation list is usually displayed on the side or the bottom, showcasing some other objects which are most relevant or similar to the current one (Figure S1). Such a recommendation list is entitled "recommended articles from TrendMD" on the Science website, "we recommend" on the PNAS website, and "customers who bought this item also bought" on the Amazon website. We extract these recommendation hyperlinks (URL links) via a self-developed python-based web crawler to construct the empirical recommendation networks (RNs).[33,34] Specifically, we adopt a breadth-first search strategy (snowball sampling) to extract empirical RNs. Firstly, a set of seeds (initial objects) are selected. For each seed, we collect all the objects in its recommendation list with hyperlinks. The recommendation lists of the newly-collected objects will be further collected. Such

a process goes on for a certain number of steps, and then we have the corresponding RN. The employed breadth-first search strategy guarantees the complete structure among the sampled nodes. Accordingly, all ICs within the node sample are identifiable. Next, we introduce more details of data collection for Science, PNAS, and Amazon RNs respectively.

Science recommendation network

The five research articles from Science Volume 369, Issue 6509 (2020) are selected to be the seeds. The breadth-first search starts from these seeds following the recommendation hyperlinks in the list entitled "recommended articles from TrendMD" (Figure S1). A recommendation list normally consists of ten relevant articles, which are not all from Science. As articles in other journals may not have recommendation lists, we only focus on the internal recommendations. There are normally up to five internal recommendations (see Figure S2 for the out-degree distribution). The search stops after 35 steps, when no further new articles can be found. Considering the later random walk experiments, we remove articles that have no out-going links. The finalized Science RN, extracted during 11-14 September, 2020, consists of 7,730 articles and 26,338 directed hyperlinks. The Science RN is relatively small in size because articles published prior to 2015 do not have recommendation lists, perhaps due to the late implementation of the recommender system.

PNAS recommendation network

We select the three articles from the Physics and Statistics section of the PNAS Volume 116, No. 12 (2019) as the seeds. Starting from these seeds, the hyperlinks in "we recommend" list (Figure S1) that point to PNAS articles are collected. The breadth-first search continues for 11 steps. In PNAS, each article has five internal recommendations. However, recommendations of new articles encountered in the last step are not taken into consideration. Therefore, after the removal of articles without out-going links, some articles' degrees are less than five (see out-degree distribution in Figure S2). Eventually, the PNAS RN consists of 59,479 articles and 261,394 recommendation hyperlinks. The PNAS RN was collected during the period of 26 March to 12 April, 2019.

Amazon recommendation network

Amazon's "customers who bought this item also bought" list is probably the most well-known RN. We first select the top-three kindle books from the bestseller list (assessed on 3 April, 2019) as the seeds. Starting from each seed, a search is performed following the recommendation hyperlinks. For each product, Amazon normally provides 100 recommendations, which are displayed on different pages. Depending on the window size of the web browser, there are normally 3 to 8 recommendations displayed on each PAGE (see Figure 1A). One needs to turn to the next PAGE to check the following recommendations. For each kindle book, we collect its top-five recommendations due to the following considerations: (i) customers would pay most attention to the recommendations on the first PAGE; [52] (ii) the computational complexity would be sharply increased if more recommendations are considered. We skip recommendation hyperlinks to other kinds of products and only collect kindle books from the recommendation lists (actually, recommendations of kindle books are mostly also kindle books). Such search continues for 16 steps. For the kindle books newly collected at the last step, we go through their full recommendation lists to find hyperlinks connecting to an existing kindle book as possible. At last, the kindle books with out-degrees of 0 or 1 are removed. The out-degree distribution of the finalized network is shown in Figure S2, where most kindle books are of out-degree 5. The Amazon RN, extracted during April 2019, consists of 119,636 kindle books and 584,093 recommendation hyperlinks.

Strategy for identification of information cocoons

According to the definition, an information cocoon (IC) is an induced subgraph of an RN that satisfies two criteria: (i) it is strongly connected; and (ii) its nodes do not have any directed link pointing to external nodes.

We first find all potential ICs by employing a breadth-first search starting from every node. Taking a small network as an example (Figure S3), if we start from node 6, the first step of the search will find node 1 and node 3, and in the second step node 2 will be found. Then the search cannot find any new nodes, namely no out-going links pointing to an unfound node. The search from node 6 returns a potential IC {6, 1, 2, 3}. Analogously, searches starting from all nodes return four potential ICs (searches may return the same

results), say $\{1,2,3\}$, $\{4,1,2,3\}$, $\{5,4,1,2,3\}$, and $\{6,1,2,3\}$. Obviously, all potential ICs satisfy the second criterion. Next, we check the first criterion. To do so, we calculate the reachability matrix $T^{N\times N}$ of the RN, which describes the transitive closure between any pair of nodes through the directed links, namely $T_{ij} = 1$ if there is a directed path from node i to node j, and $T_{ij} = 0$ otherwise. A potential IC with a node set C can be confirmed as an actual IC only if $\prod_{i \in C, j \in C} T_{ij} = 1$. To reduce the computational complexity, the potential ICs can be ranked in terms of size, and the above examination can start from the smallest ones. If a potential IC passes the above examination, all its supersets of nodes can be directly falsified, since there is no link from an IC node to an external node.

Datasets of user-object interactions

To test whether similarity-based recommendation techniques will lead to the emergence of ICs, we derive and analyze four RNs by disparate similarity-based algorithms, based on four real-world user-object interaction datasets. The Steam dataset consists of 5,094,082 records of 87,626 users purchasing 10,978 video games.[49] The Yelp dataset (downloaded from www.yelp.co.uk/dataset_challenge) consists of 1,569,264 reviews posted by 366,715 users on 60,785 local businesses (e.g., restaurants, bars, etc.). The Epinionsdataset consists of 586,359 reviews posted by 39,588 users on 61,273 products.[50] The MovieLensdataset consists of 22,884,377 ratings by 247,753 users to 33,670 movies.[51] The above four datasets are naturally described by bipartite networks. The distributions of user degrees and object degrees are reported in Figure S4, both exhibiting heavy-tailed patterns.

Extended similarity indices.

Besides the Common Neighbor index, we have also tested other three widely applied similarity indices, namely the Jaccardindex,[40,41] Salton index,[40,41] and Heat Conduction index.[16,42] Jaccard similarity calculates the similarity between two objects α and β as

$$s_{\alpha\beta}^{JAC} = \frac{\sum_{u \in U} b_{u\alpha} \cdot b_{u\beta}}{k_\alpha + k_\beta - \sum_{u \in U} b_{u\alpha} \cdot b_{u\beta}} + \epsilon,$$ (Equation 8)

where $\epsilon \rightarrow 0$ is a tiny randomnumber used to eliminate degeneracy caused by the same similarity scores, and k_α and k_β are object degrees of α and β. In the Jaccard index, the similarity between a pair of large-degree objects is depressed in comparison with the Common Neighbor index. Salton index reads

$$s_{\alpha\beta}^{SAL} = \frac{\sum_{u \in U} b_{u\alpha} \cdot b_{u\beta}}{\sqrt{k_\alpha k_\beta}} + \epsilon,$$ (Equation 9)

with a slightly different way to penalize large-scale objects. The Heat Conduction index is defined as

$$s_{\alpha\beta}^{HC} = \frac{1}{k_\beta} \sum_{u \in U} \frac{b_{u\alpha} \cdot b_{u\beta}}{k_u} + \epsilon,$$ (Equation 10)

where k_u is the user degree of u. Notice that, the Heat Conduction index is asymmetric and it penalizes not only large-degree objects, but also large-degree users.

We apply Jaccard, Salton, and Heat Conduction indices to construct RNs from the four datasets respectively with $L = 5$. The results are reported in Figure S5. Analogous to the Common Neighbor index, RNs derived by the above three similarity indices are all largely unnavigable. An exception is the Yelp RN derived by the Heat Conduction index, whose navigability ($\Omega = 2.74\%$) is significantly higher than other RNs. Yet, it is still far lower than the expected value for random RNs ($\Omega = 63.21\%$). Similar to the case for the Common Neighbor index, the low navigability of these derived RNs is resulted from ICs, which monopolize the traffic of the random walk, as reported in Table S3.

Setting and results of the online experiment

We have tested a variant of the flexible recommendation strategy on the video streaming system AiQiYi (www.iqiyi.com, NASDAQ:IQ). Our experiment was employed in the mobile application of AiQiYi (a screenshot is shown in Figure S9), which attracts on average about 1×10^8 daily active users, who open the application and have at least one clicking or scrolling action. Among these active users, about 80% will play at least one video, and about 40% of video plays are evoked by machine-generated recommendations. The

recommender system of AiQiYi is thus regarded as very influential on users' information accessing behavior. Therefore, AiQiYi is an ideal platform for our experiment.

The most important and attractive recommendation module of AiQiYi mobile application is the list of "*Guess YouLike*", displayed on the landing PAGE, as shown in Figure S9. The corresponding recommendations are generated by a sophisticated recommender system, with the fundamental structure shown in Figure S10. It employs multiple recall algorithms, each of which operates independently to identify a set of videos that are most likely to fit the target user's interests. All these selected videos will go through a grand ranking model to generate a list of relevant videos. A regulation module is then used to filter out videos with violent, porno, or brand-conflicting content. The most relevant videos that passed such regulation filtering constitute the final recommendation list, which will be displayed in "Guess YouLike".

The recall algorithms are diverse, including not only personalised algorithms like item-based and user-based collaborative filtering, but also less-personalised approaches such as the global or regional trending list to identify the most popular videos. The item-based collaborative filtering (ICF) is one of the most important recall algorithms in the AiQiYi recommender system. ICF calculates similarities among videos according to all users' co-accessing (e.g., co-clicking, co-playing, etc.) activities, and that two videos are generally more similar to each other if many users have clicked and/or played both of them. ICF will identify a set of videos that are most similar to the target user's recently clicked/played videos.

In our experiment, 5% randomly selected users (about 4×10^6 users on weekdays, and 5×10^6 users on weekends) are regarded as treatment users. For such users, we apply the flexible recommendation strategy to the ICF module. Specifically speaking, for the 5% treatment users, ICF will randomly select 5 videos from the top 10 most similar ones (corresponding to $\lambda = 2$), while for the 95% control users, ICF will directly select the top 5 most similar ones.

We use two metrics to evaluate the experimental performance. The playing rate (PR) is the number of playings on recommended videos divided by the number of clicks on these recommendations. PR describes that among all clicked recommendations, how many are actually watched. This metric is widely applied in industry since it well reflects how the recommendations fit users' true interests. On the landing PAGE (Figure S9), there is only a thumbnail poster for each recommended video, which may be not enough for the target user to decide for playing or not. If the target user clicks the recommended video, more information will be exhibited, such as the summary of the video, the producer and director of the video, the number of cumulative playings, and so on. The target user could choose to play the video or go back to the landing PAGE. This is to say, in the real scenario, clickings do not necessarily mean success, and to click but not to play usually implies not-so-good user experiences. Therefore, PR is better to measure the quality of recommendations than the clicking rate. Another metric, named playing duration (PD), is defined as the average playing duration over all played recommendations. PD can be regarded as a metric for a deeper level of user satisfaction: whether they prefer to watch for a long duration or drop out quickly. In addition, PD is commercially meaningful as longer playing corresponding to more advertisements.

The experiment was performed in AiQiYi mobile application from November 3rd to November 9th 2020, covering a full week. The daily results are reported in Table S4. The PR of the treatment group is higher than that of the control group in every day, indicating that the flexible strategy can indeed improve the fitting to users' interests. A two-sample paired *t*-test is further applied to the PR values of the treatment and control group with a null hypothesis of the equal mean. The *p*-value of such a test is 0.0082, indicating that the PR values of the two groups significantly differ from each other. However, the *p*-value of *t*-test on PD is 0.4715, suggesting that such improvement is not significant. As a conclusion, the flexible recommendation strategy can significantly promote the likelihood of users to play the recommended videos, but will not make much difference in the average duration of those plays.

The improvement in PR seems to be minor, but it largely reflects the effectiveness of the flexible recommendation strategy due to the following reasons. (i) The original AiQiYi recommender system is elaborately designed and well-trained to optimize PR and PD. Accordingly, any tiny improvement is not easy. (ii) The AiQiYi recommender system consists of a series of recall algorithms, but we only apply the flexible strategy to one of them. As there are many recall algorithms, it is not surprising that the eventual differences are

minor. With these reasons considered, the significance of PR improvement as indicated by the *t*-test demonstrates the efficacy of the proposed flexible strategy.

Quasi information cocoons

To promote the applicability of the strictly-defined IC, we relax the definition to consider the so-called quasi information cocoon (quasi-IC or QIC for short), which is a set of strongly connected nodes with very few links pointing to outside nodes. Figure S11 shows a typical example. Similar to the case of ICs, nodes belonging to QICs are called QIC nodes, links from non-IC nodes to QIC nodes are trapping links, and links from QIC nodes to non-IC nodes are escaping links.

Consider an RN $G(V, E)$, where V is a set of N nodes and E is the set of directed links which can be characterized by an adjacency matrix $A = \{a_{ij}\}$ where $a_{ij} = 1$ if node i has a directed link pointing to j and $a_{ij} = 0$ otherwise. Given a subset $C \subseteq V$, the escaping probability of the node set C can be defined as the ratio between its escaping links and total out-going links, say,

$$p^e(C) = \frac{\sum\limits_{i \in C, j \notin C} a_{ij}}{\sum\limits_{i \in C} k_i^{out}}, \qquad \text{(Equation 11)}$$

with k_i^{out} being the out-degree of node i. An induced subgraph $G[C]$ is a QIC if $G[C]$ is strongly connected and $p^e(C)$ is no larger than a predefined threshold p_c^e. Meanwhile, the trapping probability of a QIC can be defined as the ratio between its trapping links and total links originating from non-QIC nodes, that is,

$$p^t(C) = \frac{\sum\limits_{i \notin C, j \in C} a_{ij}}{\sum\limits_{i \notin C} k_i^{out}}. \qquad \text{(Equation 12)}$$

While calculating the expected navigability for network with ICs (Equation 4 and 5), we have assumed that $p^t = c/N$. This is based on the assumption that the in-degrees of nodes are uniformly distributed. However, the in-degree distributions of empirical and derived RNs are with heavy-tailed distributions. Meanwhile, ICs, or QICs, with large in-degrees would have a much stronger impact on navigability, because of the higher chance of falling into the IC or QIC. Thus, we consider also the trapping probability to reflect the varied in-degrees of QIC nodes.

The impact of QICs on navigability in random RNs can also be analyzed. Suppose in a random RN with N nodes, there is a number of QICs with in total c QIC nodes. The QICs have an average escaping probability p^e, while the summation of their trapping probabilities is p^t. A randomly surfing user has a probability of p^t to visit a QIC node at every step, and thus the expected number of steps until the user falling into the QIC is $\langle s_0 \rangle = 1/p^t$. In other words, for a user who starts the random walk from a non-QIC node, it takes $\langle s_0 \rangle$ steps to fall into the QIC. When within the QIC, the user has a probability p^e to escape at every step, and thus the expected number of steps to escape the QIC is $\langle s_c \rangle = 1/p^e$. After that, the user could fall into the QIC for the second time, third time, and so on. The number of cycles of getting out and in the QIC within N steps can be thus calculated as

$$K = \frac{N - \langle s_0 \rangle}{\langle s_c \rangle + \langle s_0 \rangle} = \frac{(N - 1/p^t)p^e p^t}{p^e + p^t}. \qquad \text{(Equation 13)}$$

The expected number of effective steps of the random walk, which is defined as the steps outside QIC, can be written as

$$\langle s_e \rangle = (K + 1)\langle s_0 \rangle = \frac{(N - 1/p^t)p^e}{p^e + p^t} + \frac{1}{p^t}. \qquad \text{(Equation 14)}$$

For such effective steps, the number of distinct nodes that could be visited is given by Equation 2. For the steps within the QIC, the number of distinct nodes that could be visited is approximately $S - 1$, where S represents the size of the visited QIC, when the escaping probability is very low. Accordingly, the expected number of distinct nodes being visited during an N-step random walk is the summation of such two parts, and the navigability is thus,

$$\Omega = \frac{n(t = \langle s_e \rangle) + S - 1}{N} = 1 - e^{-\langle s_e \rangle / N} + \frac{S - 1}{N}. \qquad \text{(Equation 15)}$$

Such a prediction is actually in line with the prediction for RNs with ICs (Equation 5). If we simplify trapping probability as $p^t = c/N$, the above equation can be updated as

$$\Omega\left(p^t = \frac{c}{N}\right) = 1 - e^{-\frac{1}{c}\left[1 + \frac{Np^e(c-1)}{Np^e+c}\right]} + \frac{S-1}{N}. \qquad \text{(Equation 16)}$$

For ICs, the escaping probability is $p^e = 0$, which gives us

$$\Omega\left(p^t = \frac{c}{N}, p^e = 0\right) = 1 - e^{-\frac{1}{c}} + \frac{S-1}{N}, \qquad \text{(Equation 17)}$$

reproducing Equation 5.

To validate such prediction, we simulate random walks in random RNs with QICs, the results of which are reported in Figure S13. In general, QICs with low escaping probabilities and high trapping probabilities have stronger impacts on navigability. The simulation results well follow the predictions as specified by Equation 15 especially when the trapping probability is high and escaping probability is low. When the QIC has a trapping probability $p^t = 0$, the RN can be regarded as IC-free (also QIC-free), and thus the navigability is as high as 58.71%, regardless of the escaping probability. For any non-zero trapping probabilities, a QIC with low escaping probability has a severe impact on navigability. On the other hand, the increase in the trapping probability would also make the RN unnavigable.