User tag weight assessment based on fuzzy theory in mobile social networks

Min Gao
Fujian Normal University

Li Xu (xuli@fjnu.edu.cn)
Fujian Normal University https://orcid.org/0000-0002-8972-3373

Xiaoding Wang
Fujian Normal University

Xinxin Zhang
Fujian Normal University

Research

Keywords: User Tag Weight, User Profiling, Mobile Social Networks, Fuzzy Theory, Genetic Algorithm

Posted Date: January 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-52476/v2

RESEARCH

User tag weight assessment based on fuzzy theory in mobile social networks

Min Gao[1], Li Xu[2*], Xiaoding Wang[1] and Xinxin Zhang[1]

*Correspondence: xuli@fjnu.edu.cn
[2]Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou, Fujian, China
Full list of author information is available at the end of the article

Abstract

Mobile social network supports mobile communication and asynchronous social networking. For enterprises, how to provide better services and create greater business value through the data and information provided by users is crucial. For example, enterprises need to build user profiles to achieve personalized recommendation and precision marketing. In view of the data modeling stage of user profile, we propose a method to evaluate user tag weight, which includes two steps. Specifically, we introduce fuzzy theory to get the initial weight interval. Then, genetic algorithm with single point crossover is used to optimize user tag weight. Experiment results show that the proposed method has better performance than other three methods applied to recommendation system.

Keywords: User Tag Weight; User Profiling; Mobile Social Networks; Fuzzy Theory; Genetic Algorithm

1 Introduction

Mobile social networks (MSNs) [1-7] are networks with device mobility and social communication. With phones or tablets, people who share common interests can create a profile, multimedia posts, instant messaging and play social gaming. With the development of a variety of online social platform, these platforms (such as QQ, weibo, circle of friends, etc.) are much more than a sort of social tool for user to communicate, they are the main medium for the generation and dissemination of social information. Hence, the mobile communication network layer includes network communications among mobile devices, and the user social relation layer contains social interaction among users. Figure 1 is the MSNs model of user social relation layer. The massive data carried by mobile social networks contains massive commercial value. However, user information is complex, and there are cases of missing or false information [8]. Under the above background, user profiling aims to build a quantifiable information representation, to describe user characteristics and to mine user relations.

User profile refers to a tagged user model abstracted based on user's basic attributes, user preferences, living habits, user behaviors and other information. Each tag and tag weight is a vector of the user, and a user can be understood as the sum of multiple vectors (tags) of the hyper dimensional space. Users described by data can finally be recognized by the computer, and user profile application can be realized on this basis. The determination of tag weight has a great impact on subsequent user profiling based recommendations and precision marketing. Existing tag weight algorithms include PageRank algorithm [9], TF-IDF algorithm [10], BM25 algorithm [11], etc. While these methods have some drawbacks. PageRank algorithm fails to quickly improve the score of new high-quality pages, TF-IDF algorithm uses term frequency to measure the importance of a keyword, which is not

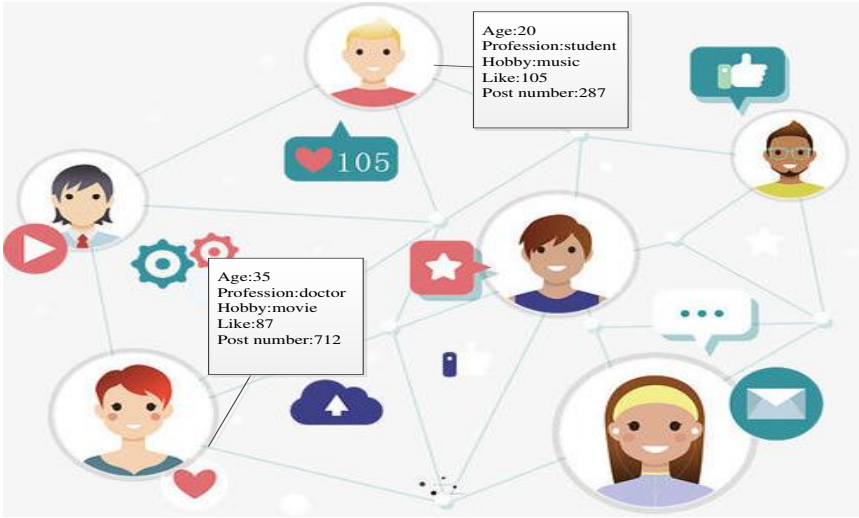

Figure 1. The MSNs model of user social relation layer.

comprehensive enough, and BM25 algorithm does not take the relevant features of the keyword itself into account.

Inspired by the above problems, we introduce fuzzy theory [12-14] and genetic algorithm [15-18] to explore this problem, where the former is used for the determination of the weight interval, and the latter performs the optimal weights. Fuzzy theory is the study of many boundary unclear things in life as a theoretical tool, because of the differences between people, there are certain subjective judgment of fuzzy things. Fuzzy theory can be used to evaluate things reasonably, and it's role is to get the preliminary fuzzy delimitation. And genetic algorithm (GA) is a random adaptive global search algorithm and it has good global searching ability, which can search out all the solutions in the solution space quickly without falling into the trap of local optimal solution. The results of the experiment show that this work has higher precision compared with other three algorithms.

In this work, we focus on the problem of evaluation of user tag weight, which is essential for user profiling and recommendation system. We design membership function to obtain the searching space, and we use genetic algorithm to get the optimized user tag weight. Totally, our contributions are as follows.

- We transform the problem of user tag weight evaluation into an optimal solution seeking problem. Then we design membership degree function to get the fuzzy boundaries of all user tag weights, and we use genetic algorithm to get the optimal solution of each user tag weight.
- We start from three dimensions, (i.e., basic tag, network tag, and behavior tag) and give a division of different kinds of user tags, which complies with the background of mobile social network.
- We conduct experiments on datasets crawled from two social platforms and the results outperform existing methods on three evaluation metrics.

The rest of this paper is organized as follows. Related works are reviewed in section 2. The proposed method is introduced in Section 3, and Section 4 is the part of experiment and analysis. Finally, section 6 draws conclusion.

2 Related work

Data mining on MSNs has been widely concerned by academia and industry. In opportunistic mobile networks, Zhou *et al.* [19], considering the freshness of the content and the transmission cost from the cellular network to the initial seeds, presented two methods for seed selection to find the optimal number of initial seeds and maximize the overall content utility value of nodes in the network in order to solve the problem of utility optimization. They also made predictions on nodes' social contact patterns with a temporal perspective, and designed TCCB to improve the performance of data forwarding in [20]. With regard to user profiling based on mobile data, there are three main methods in the stage of data modeling: PageRank algorithm, Term Frequency-Inverse Document Frequency (TF-IDF) algorithm, and BM25 algorithm. The PageRank algorithm was developed by Google's founders, and it's the core algorithm of Google's search engine. It can be used to identify the importance of a web page [9], which can also be used to calculate keyword weights. The PageRank algorithm gives each page an importance rating, which is a recursively defined metric, if an important page links to it, then the page becomes important. This definition is recursive, because the importance of a page references the importance of other pages that link to it. while the TF-IDF algorithm is a classical algorithm for calculating keyword weights, which is a numerical and statistical metric for evaluating the importance of keywords in relation to the text [10]. The more frequently a keyword appears in the text, the more important it is, but at the same time, its importance decreases the more frequently it appears in the entire text set. And the advantages of TF-IDF algorithm are simple and fast, and the results are in line with the actual situation. The disadvantage is that simply using "word frequency" to measure the importance of a word is not comprehensive enough, and sometimes important words may not appear many times. Moreover, this algorithm fails to reflect the position information of the word, and the word near the front of the position is regarded as the same importance as the word near the back of the position, which is incorrect. As for BM25 algorithm was originally applied to information retrieval, often used to retrieve relevance scores [11]. The BM in BM25 means best matching, and BM25 is considered to be the most advanced TF-IDF class search function used in document retrieval. The algorithm is mainly used in information retrieval, which core idea is that parsing of the query to generate the semantics, and then obtaining the result text D for each retrieval and calculating the correlation between each semantics and the text. Finally, calculating the correlation between each semantics and the text by weighted summation, and then the score of to the query and the correlation score of the text.

Different from the above three methods, we have realized conversion from the evaluation of user tag weight to the optimization seeking. We address this problem in two steps. For the first time, we try to introduce fuzzy theory and combine global optimization algorithm (GA) to solve the problem.

3 Method

3.1 Problem description

The key to build user profiling is data modeling, i.e., determining user tag weight. Inevitably, there is subjectivity in determining the weight, so we combine fuzzy theory and genetic algorithm to design an optimization method to apply it to the evaluation of user tag weight. As shown in Figure 2, the proposed method includes four steps: selection of user tag, determination of membership degree function, determination of tag weight interval,

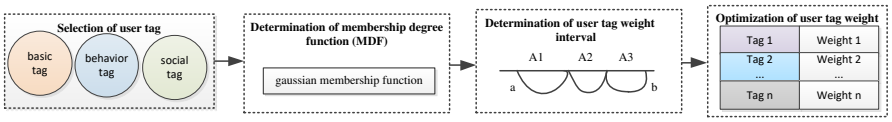

Figure 2. Flow chart of the proposed method.

Table 1 Some notations used in this article.

Symbol	Description
Opi	ith opotimization
Op-average	the mean of the previous optimizations
$Iter_{max}$	the largest number of iterations
x_i	the user tag value of the ith node
y_i	the normalized value of user tag of the ith user
$A1, A2, A3$	fuzzy set
c_i	the mean of the user tag weights
σ	the variance of the user tag weights
μ_{A_i}	the corresponding gaussian function
Y	a set of generated evaluation results
Z	random fractional matrix of each user tag
X	any set of user tag weights
x_1^i, x_2^i	the upper and lower bounds of the weight fluctuation

and optimization of user tag weight. And some notations used in this article is shown in Table 1.

3.2 Selection of use tag

Based on the characteristics of MSN, user tags are divided into three dimensions: basic tag, behavior tag, and network tag. Among them, basic tag includes some basic user attributes, such as age, gender and so on, bahavior tag involves some behaviors of users. From the perspective of users' social network, social tag is important, such as the number of neighbors and the location of users. Table 2 lists the corresponding labels and categories. Table 3 shows the initial tag weights given artificially.

3.3 Determination of membership degree function (MDF)

Firstly, the fuzzy sets [21-22] $A1$, $A2$ and $A3$ are taken to represent three levels of user tag weights, namely "small, moderate and large" respectively, and the corresponding ones are

Table 2 The information of user tag.

Basic Tag	gender	age	career	major	hobby
Behavior Tag	response	follow	post	like	repost
Social Tag	DC	CC	katz	BC	EC

Table 3 The initial tag weights.

Basic Tag	gender	age	career	major	hobby
Weight	0.1	0.2	0.6	0.4	0.8
Behavior Tag	response	follow	post	like	repost
Weight	0.5	0.1	0.7	0.85	0.2
Social Tag	DC	CC	katz	BC	EC
Weight	0.26	0.09	0.05	0.09	0.07

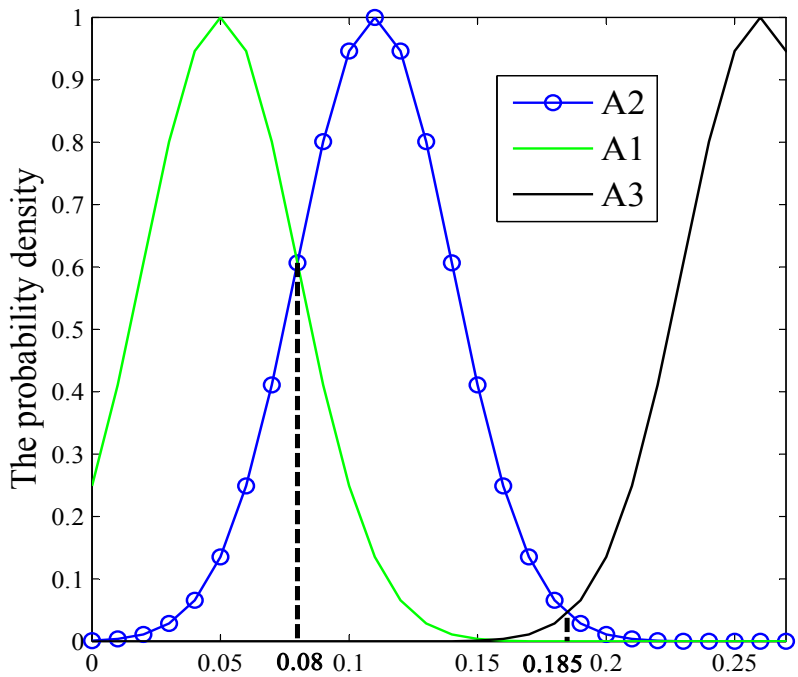

Figure 4. Guassian MDF on social tag weight.

generated MDF [23], as shown in Figure 4 (taking social tag as an example). In this paper, gaussian function [24] is used to represent fuzzy sets.

$$\mu_{A_i} = \mathrm{e}^{-\frac{(x-c_i)^2}{2\sigma^2}} \quad , x \in (0,1), i = 1,2,3 \tag{1}$$

where c_i is the mean of the user tag weights, and σ is corresponding variance.

In terms of parameter setting, in order to subdivide social tag weights, the variance of the normal MDF is determined by the interval range formed by the initial weight value. In other words, by constantly adjusting the variance, the intercept of the gaussian MDF of the fuzzy set $A2$ on the X-axis is exactly equal to the interval formed by the initial weight value. At the same time, in general, three MDF in this paper have the same variance. In terms of setting the mean value, the mean value of the three normal MDF is set as the minimum value, the mean value and the maximum value of the initial user tag weight set,

Table 4 The parameters of MDF.

Parameter setting	σ	c_i
$A1$	0.03	0.05
$A2$	0.03	0.11
$A3$	0.03	0.26

Table 5 The change interval of each user tag.

Initial weight	Fuzzy set	Interval range
0.26	$A3$	(0.185,0.26]
0.09	$A2$	(0.08,0.185)
0.05	$A1$	(0,0.08)
0.09	$A2$	(0.08,0.185)
0.07	$A1$	(0,0.08)

so as to cover the social tag weight more evenly by determining the position of MDF, relevant parameters are shown in Table 4.

3.4 Determination of user tag weight interval

The initial value of each social user weight is substituted into Equation (1) to calculate the membership degree (MD). According to the principle of maximum MD, the grade of 9 initial user tag weights is determined. The purpose of this paper is to ensure that the change of user tag weight does not exceed its existing level. Through Equation (1), the corresponding weights of the X-coordinate of the intersection of the three MDF in Figure 4 (taking social tag as an example) can be calculated as 0.08 and 0.185, respectively. Thus, the change interval of each social tag weight can be obtained, as shown in Table 5.

3.5 Optimization of user tag weight

In this paper, the weights of a group of evaluation user tags are taken as the calculation unit, and the variance of a group of calculated evaluation results thereby is taken as the fitness function, we use this fitness function to design the GA [37] to solve the following mathematical problems:

$$\max F = D(Y) \tag{2}$$

$$Y = Z \cdot X \tag{3}$$

$$x_1^i \leqslant x_i \leqslant x_2^i, i \in 1, 2, \cdots, 10 \tag{4}$$

$$\sum_{i=1}^{n} x_i = 1 \quad , i \in 1, 2, \cdots, 10 \tag{5}$$

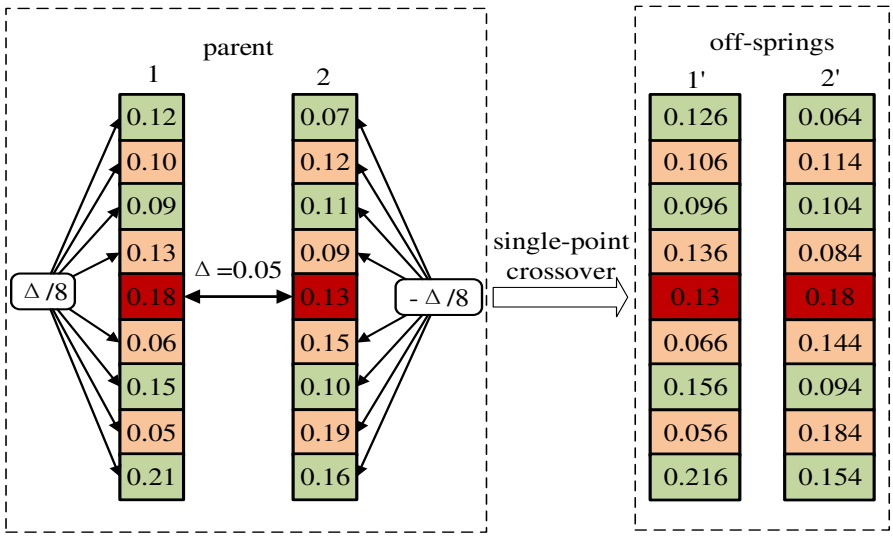

Figure 5. An example of single point crossover.

Algorithm 1 Optimization of user tag weight with GA

Input: The interval of user tag weights
Output: The set of optimization of user tag weights
Encode the indicidual of the interval of user tag;
Initialize the first population in the given interval of user tag weights;
Evaluate the variance as fitness function;
while (The overall fitness changes much || the number of iterations is less than $Iter_{max}$) **do**
  for each set of user tag weights **do**
    Perform the single point crossover to generate offspring;
    Calculate fitness for these offspring;
  end for
end while
return The optimal set of user tag weights with maximum variance

where Y represents a set of generated evaluation results, Z is the random fractional matrix of each user tag, X refers to any set of user tag weights, x_i indicates the weight of the i-th user tag in this group, and x_1^i and x_2^i respectively represent the upper and lower bounds of the weight fluctuation. Maximizing the variance of a group of evaluation results can be obtained by Equation (2). Equation (3) is the matrix representation form generated by the evaluation results. Equation (4) indicates that all weight values shall not exceed their corresponding fluctuation range. The sum of the user tag weights in the same group is 1, which can be guaranteed by Equation (5).

In the design of GA, a chromosome is composed of a set of user tag weights, so the main problem is to make sure that even after mutation and crossover, offspring chromosomes still meet Equation (5), where the sum of the weights equals 1. Here, we choose a single point crossover method to address this problem: first of all, the sum of genes will change if the two parent chromosomes in any genetic crossover occurs. And this part of changes will be shared by the other 8 genes, which ensures that the offspring chromosomes genes sum to 1, as shown in Figure 5. And if the crossover process makes the sum of genes of chromosome 1 increase Δ, we will make the other 8 genes as well as reduce $\Delta/8$ simultaneously. This method can achieve the purpose of genetic crossover and it guarantees the diversity of the population. At the same time, this method can make the changes of uncrossed genes exceed

Table 6 The parameter setting of simulation.

Population size	$Iter_{max}$	Crossover probability	Mutation probability
100	1000	0.9	0.01

Table 7 The matrix of random scoring on user tag.

Basic Tag	gender	age	career	major	hobby
score	5	4	2	3	1
Behavior Tag	response	follow	post	like	repost
score	3	5	2	1	4
Social Tag	DC	CC	katz	BC	EC
score	1	4	3	2	5

the fluctuation range as less as possible. What's more, this method will not cause changes in the sum of genes. The optimization of user tag weight can be conducted by randomly generated data and weight interval obtained by fuzzy theory and GA. We set the algorithm to run 6 times, and each time is conducted based on different random score within the user tag weight interval. The relevant parameters of the simulation are shown in Table 6.

3.6 Case study

Here we use the matrix of normalized X_{3*5} as the set of user tag weights from Table 3. And a random scoring matrix Z_{5*3} of user tag from Table 7. With Equation (3), we can get the matrix of Y_{5*5}, then we take the result of Equation (4) as the fitness function.

According to Algorithm 1, the first step is to encode users' tag weights. Here we use float encoding with fast convergence and high optimization accuracy (line 1). Then we need to initialize the first population according to the interval of user tag weights (line 2). Then, the iterative optimization phase begins (line 4–line 9). And the convergence condition is that there is no large change in the fitness function or the number of maximum iterations $Iter_{max}$ has been reached (line 4). In the process of iteration, we choose a simple single point crossover operation without considering the mutation process. This is because the crossover operator is used as the primary operator in the genetic algorithm because of its global search capability, while the mutation operator is used as the secondary operator because of its local search capability. The goal of this paper is to verify the feasibility and the effectiveness of the designed method after we have transformed the problem to an optimal problem. At the end of the iteration, we can obtain the sought optimal solution (line 10) in Table 8.

4 Experiment and analysis

We introduce two real-world datasets and selecte some evaluation metrics to quantitatively evaluate our proposed method. We use two classic methods introduced in related work as comparisons. Relevant implement details are also shown in this section.

Table 8 The matrix of optimal solution of user tag weights.

1	0.07	0.072	0.073	0.071	0.074	0.073	0.072	0.0725
2	0.09	0.089	0.086	0.088	0.089	0.092	0.093	0.0895
3	0.09	0.092	0.091	0.09	0.089	0.09	0.088	0.09
4	0.26	0.255	0.253	0.254	0.265	0.267	0.255	0.258
5	0.05	0.052	0.051	0.048	0.049	0.05	0.049	0.0498

Table 9 The information of datasets.

Number of posted messages	Time span	Average text length
2000	two months	35
15000	eight months	500

4.1 Dataset information

Two datasets in this paper are derived from our crawling of information on two real social networks (Zhihu and Sina Weibo). The pre-processing includes two stages of word separation and keyword extraction. The description of the pre-processed datasets is shown in Table 9.

4.2 The comparison of three evaluation metrics

In order to verify the effectiveness of our method, we compare with the existing several pop-ular methods on three evaluation metrics. **Precision** [25] is used to measure the proportion of the predicted relevant user tag samples among the actual relevant user tag samples, i.e., it indicates the accuracy of prediction. **Recall** [25] is used to measure the proportional re-lationship between the predicted relevant keywords and the relevant content of all samples, i.e., it indicates the comprehensiveness of the prediction results. **F1-score** [25] is calculated by combining the results of precision rate and recall rate, which is an evaluation criterion for the comprehensive performance of the algorithm. Only when the precision and recall rates are both high, the F1-score can be taken as the ideal value. The results are shown in Figure 6-11. The results of comparison suggest that the social network of Weibo focuses more on the breadth of content, while Zhihu focuses more on the depth of content.

We compare the performance (including prediction, recall and F1-score) of all the methods on these two datasets. From the results, we have several interesting observations and insights:

- In the experiments on dataset of Zhihu, the proposed method has higher performance on the user tag of career, while the other two methods have higher performance on the user tag of age. It is undeniable that due to the heterogeneity of individual users, a user's profession is often more important for us to build user profiling in the platform of Zhihu. It demonstrates that our method can perform better when we need to evaluate the weights of several user tags.
- In view of the experiments on dataset of weibo, the proposed method performs better on the user tag of hobby, which is consistent with the fact that hobby is more important for user profiling on Weibo. It suggests that our method can evaluate the importance of user tag flexibly.

5 Conclusion and Discussion

5.1 Conclusion

Evaluation of user tag weights is an important step in user profiling data modeling. In this work, we have tried to evaluate the weights of user tags by converting it into an optimal solution seeking problem. Specifically, fuzzy theory is used to get the interval of user tag weights, and the GA focuses on the optimal solutions. The experiment results show that the proposed method has higher precision in most cases. Because of the uncertainty of fuzzy theory and optimality seeking of GA, the proposed method can get a better performance than other methods.

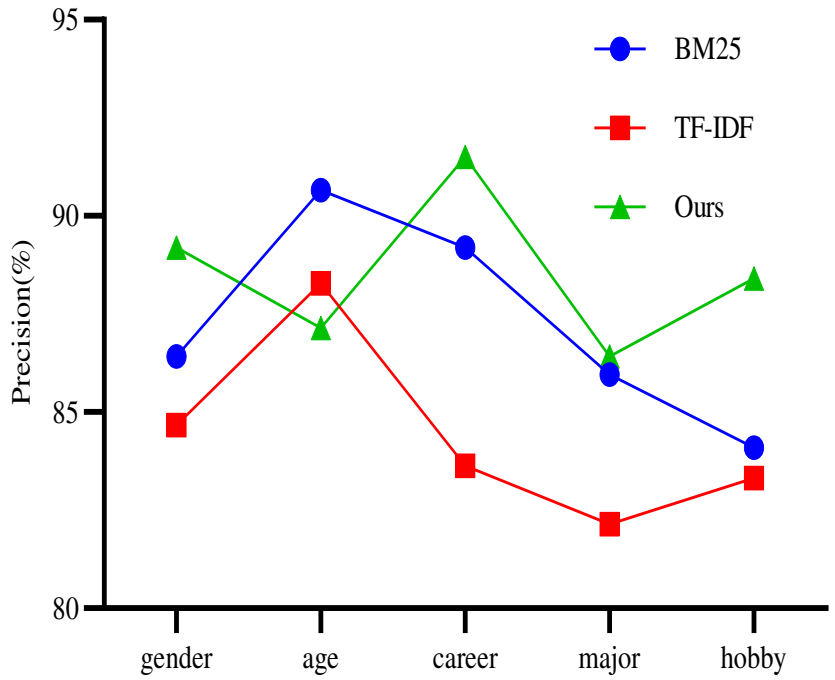

Figure 6. Accuracy comparison on dataset from Zhihu.

5.2 Discussion

The accuracy of the evaluation of user tag weight largely affects the overall perception of individual users by the service provider. In view of the evaluation of user tag weight, we have a further understanding of individuals in mobile social networks, and at the same time, users could get the improved efficiency of communication and services. But the GA used in the proposed method, whose complexity depends on the genetic operator, has uncertainty. This is where we need to think properly and try to improve in the future.

Abbreviations

MSNs: Mobile social networks; DC: Degree centrality; CC: Closeness centrality; BC: Betweenness centrality; EC: Eigenvector centrality; MDF: Membership degree function; MD: Membership degree; GA: Genetic algorithm; TF-IDF: Term frequency-inverse document frequency.

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

The authors would like to thank the National Natural Science Foundation of China (No.U1905211, No.61771140, No.61702100, No.61702103), Fok Ying Tung Education Foundation (No.171061).

Author's contributions

MG and LX conceptualized the idea and designed the experiments. MG contributed in writing and draft preparation, and XdW and XxZ supervised the research. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the National Natural Science Foundation of China (No.U1905211, No.61771140, No.61702100, No.61702103), Fok Ying Tung Education Foundation (No.171061).

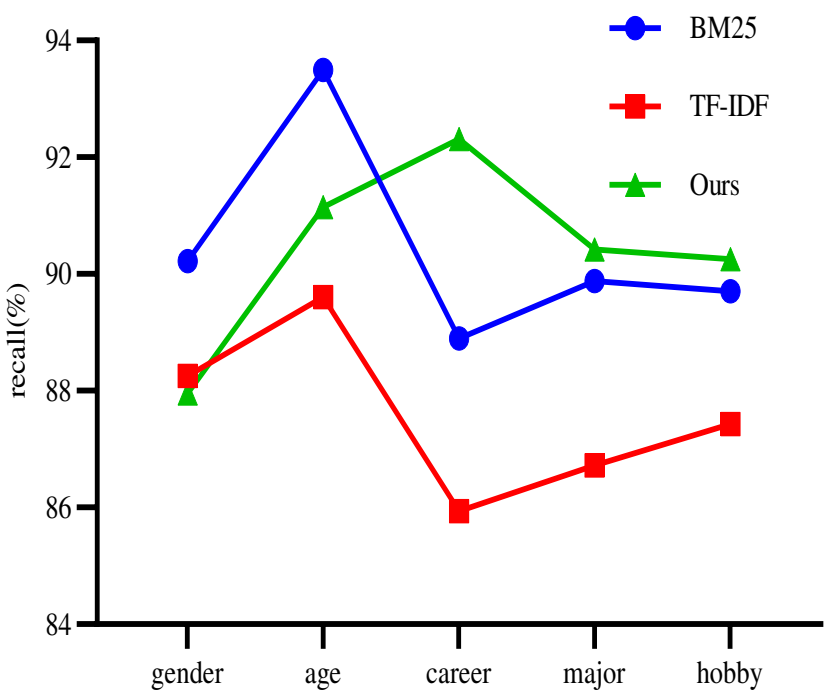

Figure 7. Recall comparison on dataset from Zhihu.

Author details

[1]College of Mathematics and Informatics, Fujian Normal University, Fuzhou, Fujian, China. [2]Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou, Fujian, China.

References

1. M. Li, Z. Jiang, W. Xie, S. Micciche, M. Tumminello, W. Zhou, R. Mantegna, A comparative analysis of the statistical properties of large mobile phone calling networks. Scientific reports. **4**, (2014)
2. X. Hu, T. Chu, E. Ngai, P. Kruchten, H. Chan, A Survey on Mobile Social Networks: Applications, Platforms, System Architectures, and Future Research Directions. IEEE Communications Surveys & Tutorials. **17**, (2014)
3. T. Qiu, B. Chen, A. Kumar, J. Ma, R. Huang, A Survey of mobile social networks: Applications, Social Characteristics, and Challenges. IEEE Systems Journal. **12**(4), 1–16 (2017)
4. C. Wei, J. Wu, C. Tan, Friendship-based location privacy in Mobile Social Networks. Int. J. of Security and Networks. **6**, 226–236 (2011)
5. A. Rizwan, S. Yan, Z. Zhu, K. Zuhaib,M. Imran, R. Saeed, A. saeed, content distribution and protocol design issue for mobile social networks: a survey. EURASIP Journal on Wireless Communications and Networking. **2019**, (2019)
6. K. Chen, H. Poor, R. Prasad, Mobile social networks. IEEE Wireless Communications. **21**, 8–9 (2014)
7. L. Humphreys, H. Hardeman, Mobiles in public: Social interaction in a smartphone era. Mobile Media & Communication. (8), 205015792092706 (2020)
8. B. Ghanem, P. Rosso , F. Rangel, An emotional Analysis of False Information in Social Media and News Articles. ACM Transactions on Internet Technology. **20**(2), 1–18 (2020)
9. L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: bringing order to the web. Technical report. (1919)
10. J. Chen, C. Chen, Y. Liang, In 2016 2nd International Conference on Artificial Intelligence and Industrial Engineering (AIIE). Optimized tf-idf algorithm with the adaptive weight of position of word, (Beijing, 2016), pp. 114—117. https://doi.org/10.2991/aiie-16.2016.28
11. Y. Lv, C. Zhai, in 20th ACM Conference on Information and Knowledge Management (CIKM). Adaptive term frequency normalization for BM25. (Glasgow, 2011), pp. 1985—1988.
12. T. Watanabe, T. Yoshikawa,T. Ito, Y. Miwa, T. Shibata, S. Saga, An infinite-resolution grid snapping technique based on fuzzy theory. Applied Soft Computing. **89**, 106112 (2020)
13. Y. Yuan, H. Chai, Knowledge fusion method based on fuzzy set theory. J. Intell. Fuzzy Syst. **38**(4), 3971–3979 (2020)
14. E. Hancer, New filter approaches for feature selection using differential evolution and fuzzy rough set theory. Neural Computing and Applications. **32**, (2020)
15. G. DAngelo, F. Palmieri, GGA: A modified genetic algorithm with gradient-based local search for solving constrained optimization problems. Inf. Sci. **547**, 136—162 (2021)

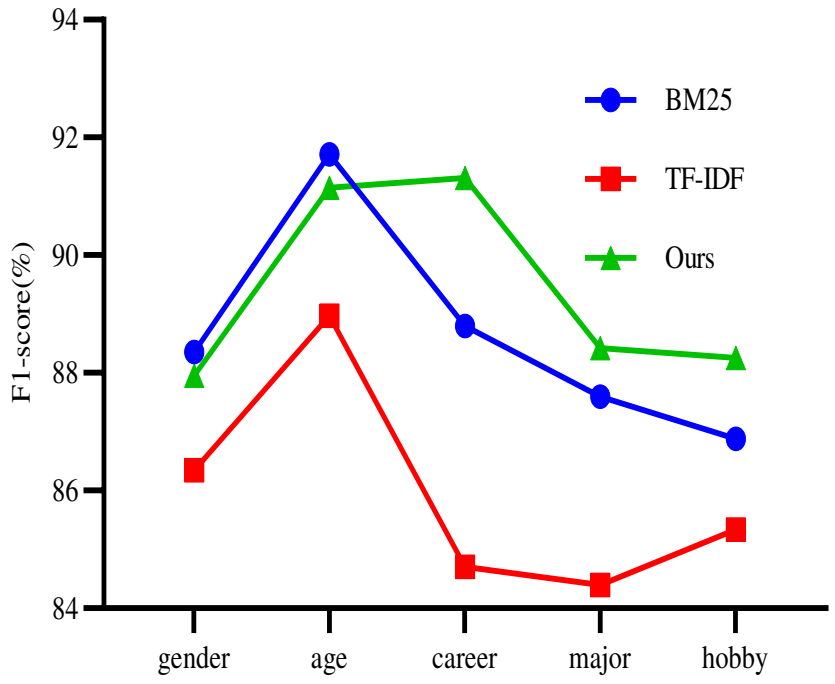

Figure 8. F1-score comparison on dataset from Zhihu.

16. X. Yue, H. Zhang, Modified hybrid bat algorithm with genetic crossover operation and smart inertia weight for multilevel image segmentation. Appl. Soft Comput. **90**, 106157 (2020)
17. P. Thanh, B. Huynh, L. Bui,New mechanism of combination crossover operators in genetic algorithm for solving the traveling salesman problem. (2020)
18. J. Chen, F. Zhao, Y. Sun, Y. Yin, Improved xgboost model based on genetic algorithm. Int. J. Comput. Appl. Technol. **62**(3), 240—245 (2020)
19. H. Zhou, V. Leung, C. Zhu, S. Xu, J. Fan, Predicting temporal social contact patterns for data forwarding in opportunistic mobile networks. IEEE Trans. Veh. Technol. **66**(11), 10372—10383 (2017)
20. H. Zhou, H. Wang, C. Zhu, V. Leung, In 2018 IEEE Wireless Communications and Networking Conference (WCNC). Freshness-aware initial seed selection for traffic offloading through opportunistic mobile networks. (Barcelona, 2018), pp. 1—6.
21. J. McCulloch, C. Wagner, On the choice of similarity measures for type-2 fuzzy sets. Inf. Sci. **510**, 135—154 (2020)
22. W. Chen,Y. Zou, Group decision making under generalized fuzzy soft sets and limited cognition of decision makers. Eng. Appl. of AI. **87**, (2020)
23. E. Kayacan, A. Sarabakha,S. Coupland, R. John, M. Khanesar, Type-2 fuzzy elliptic membership functions for modeling uncertainty. Eng. Appl. of AI. **70**, 170-–183 (2018)
24. Y. Ma, Y. Xu, Computing integrals involved the gaussian function with a small standard deviation. Journal of Scientific Computing. **78**, (2019)
25. J. Qiu, J. Tang, H. Ma, Y. Dong, K. Wang, J. Tang, In the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data 300 Mining (KDD). Deepinf: Social influence prediction with deep learning. (London, 2018), pp. 2110—2119. https://doi.org/10.1145/3219819.3220077

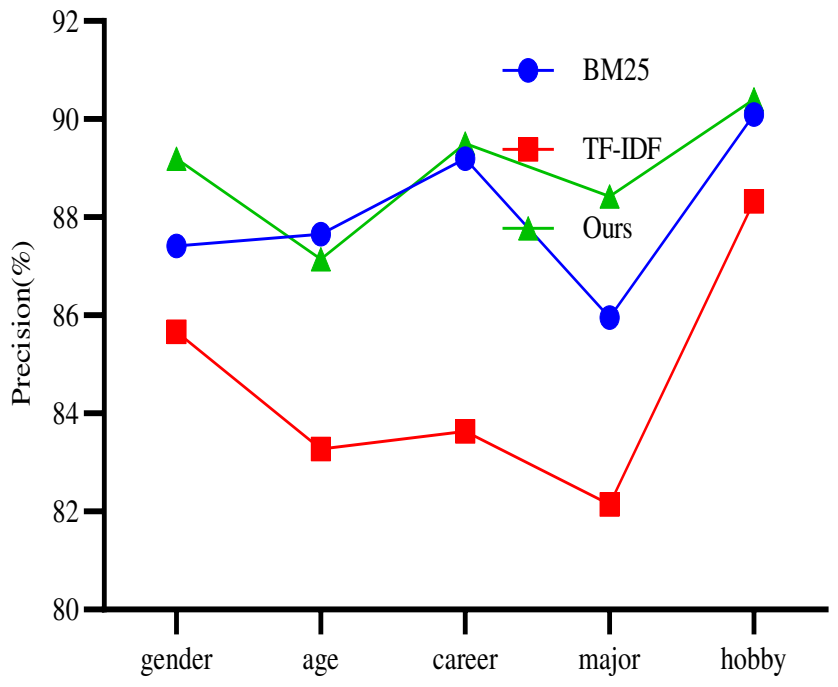

Figure 9. Accuracy comparison on dataset from Weibo.

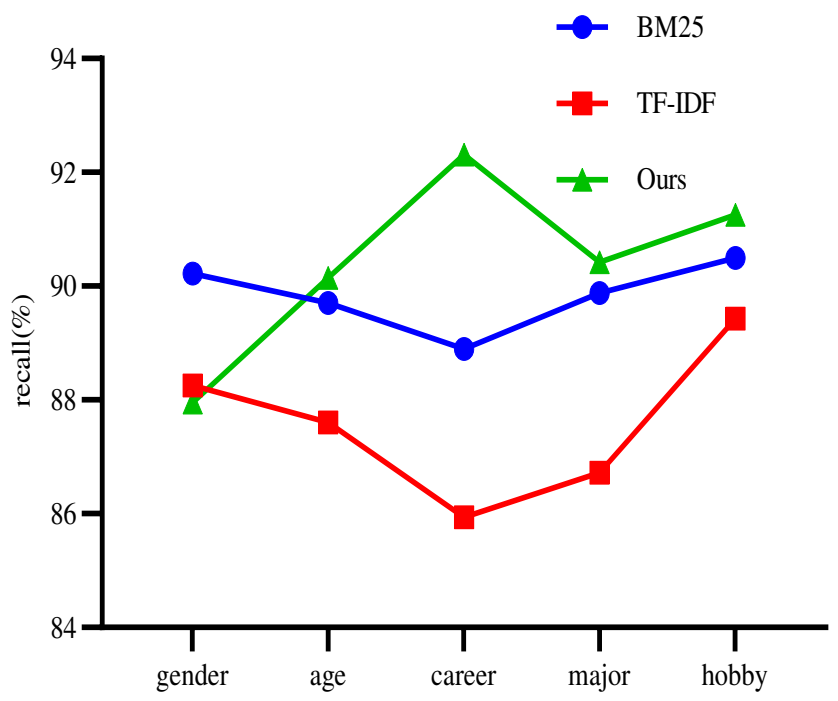

Figure 10. Recall comparison on dataset from Weibo.

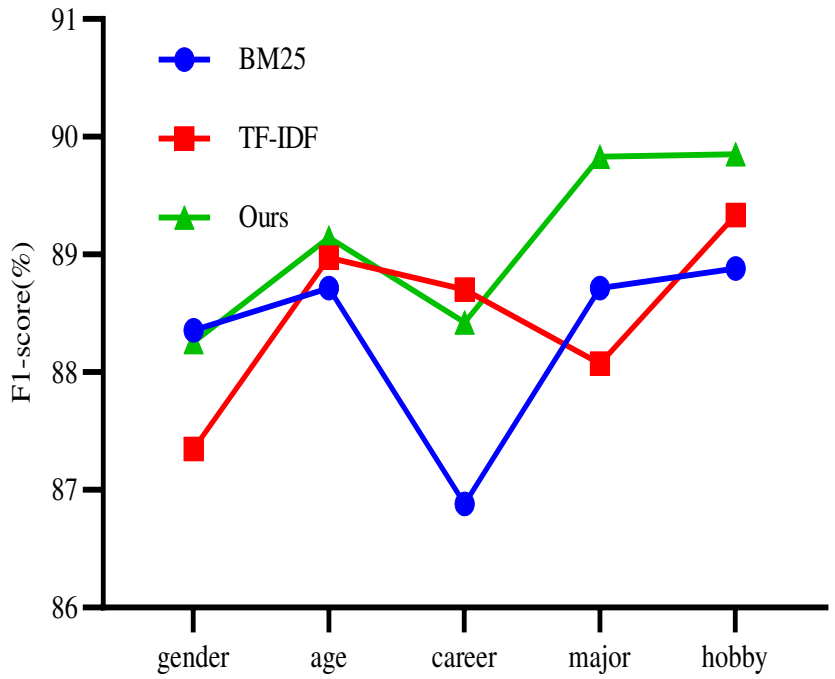

Figure 11. F1-score comparison on dataset from Weibo.

Figures

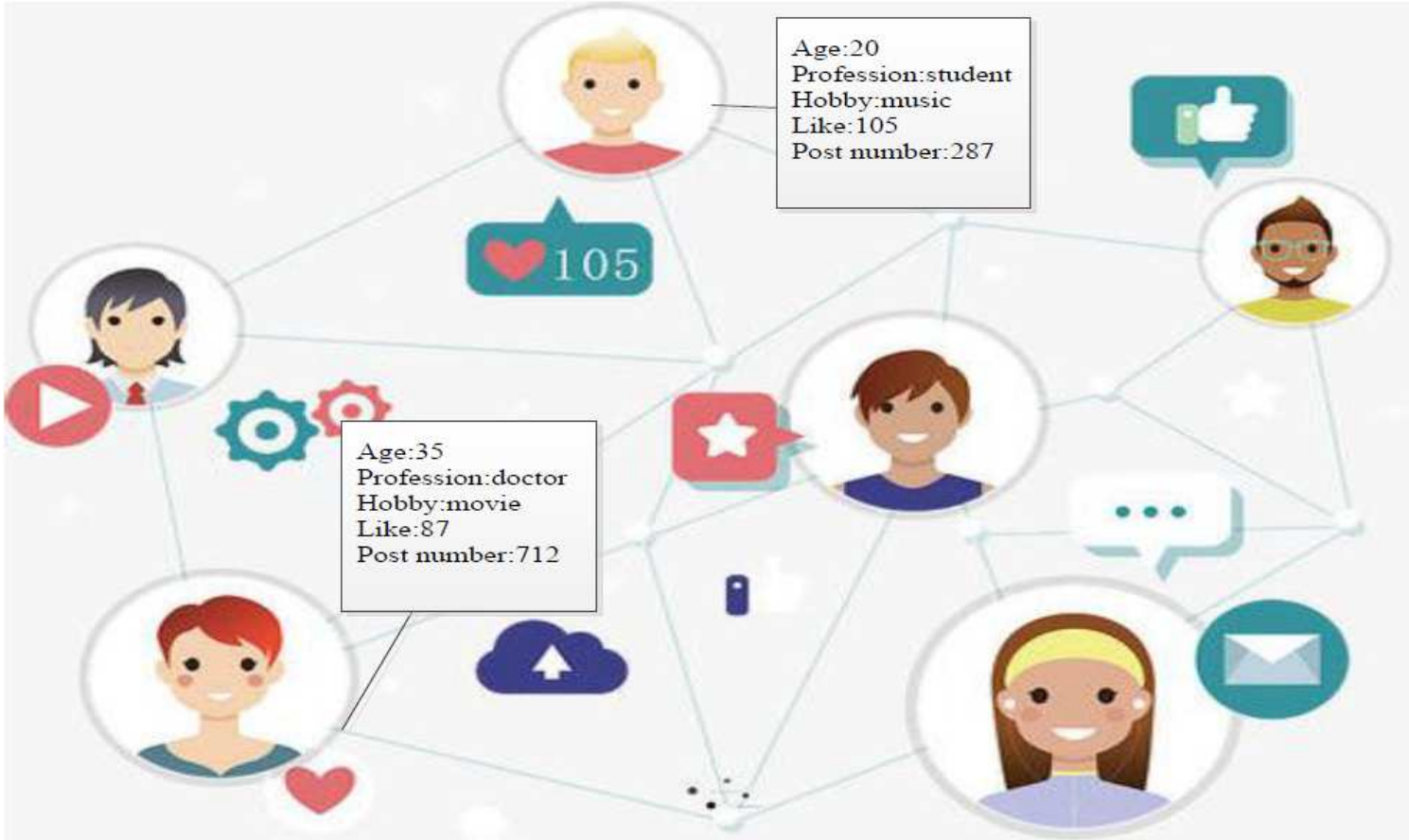

Figure 1

The MSNs model of user social relation layer.

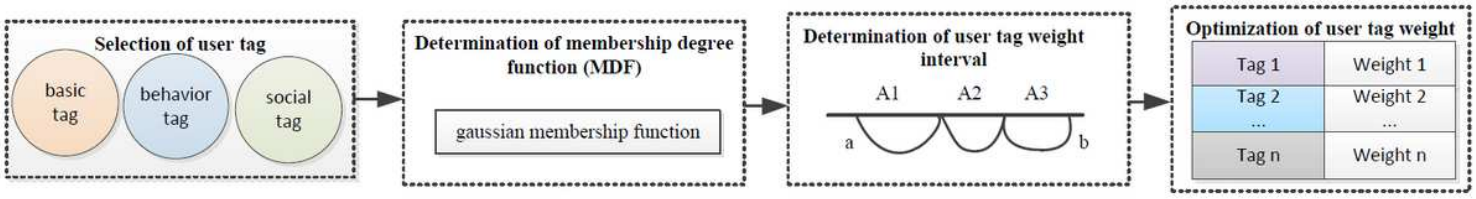

Figure 2

Flow chart of the proposed method.

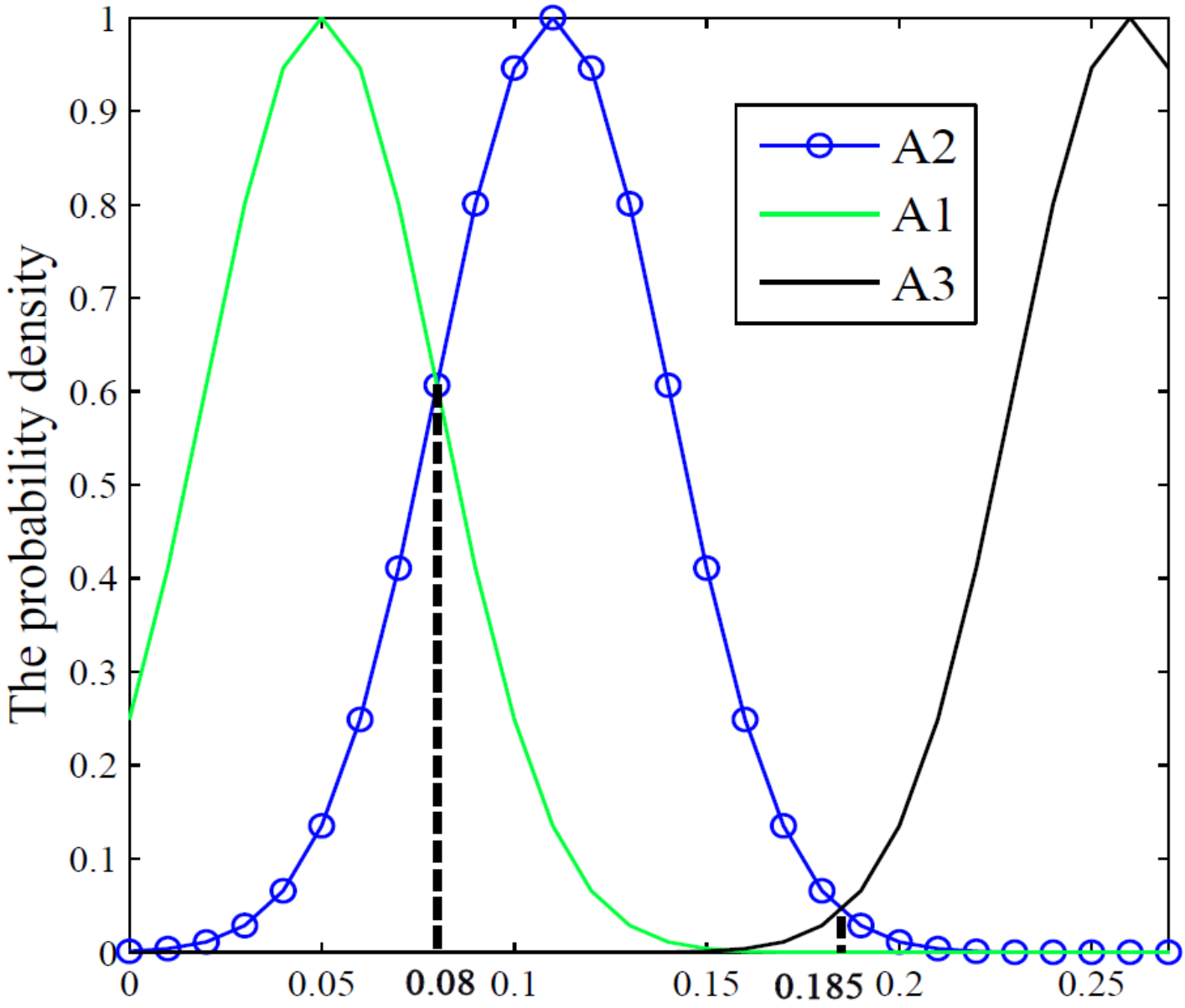

Figure 3

Guassian MDF on social tag weight.

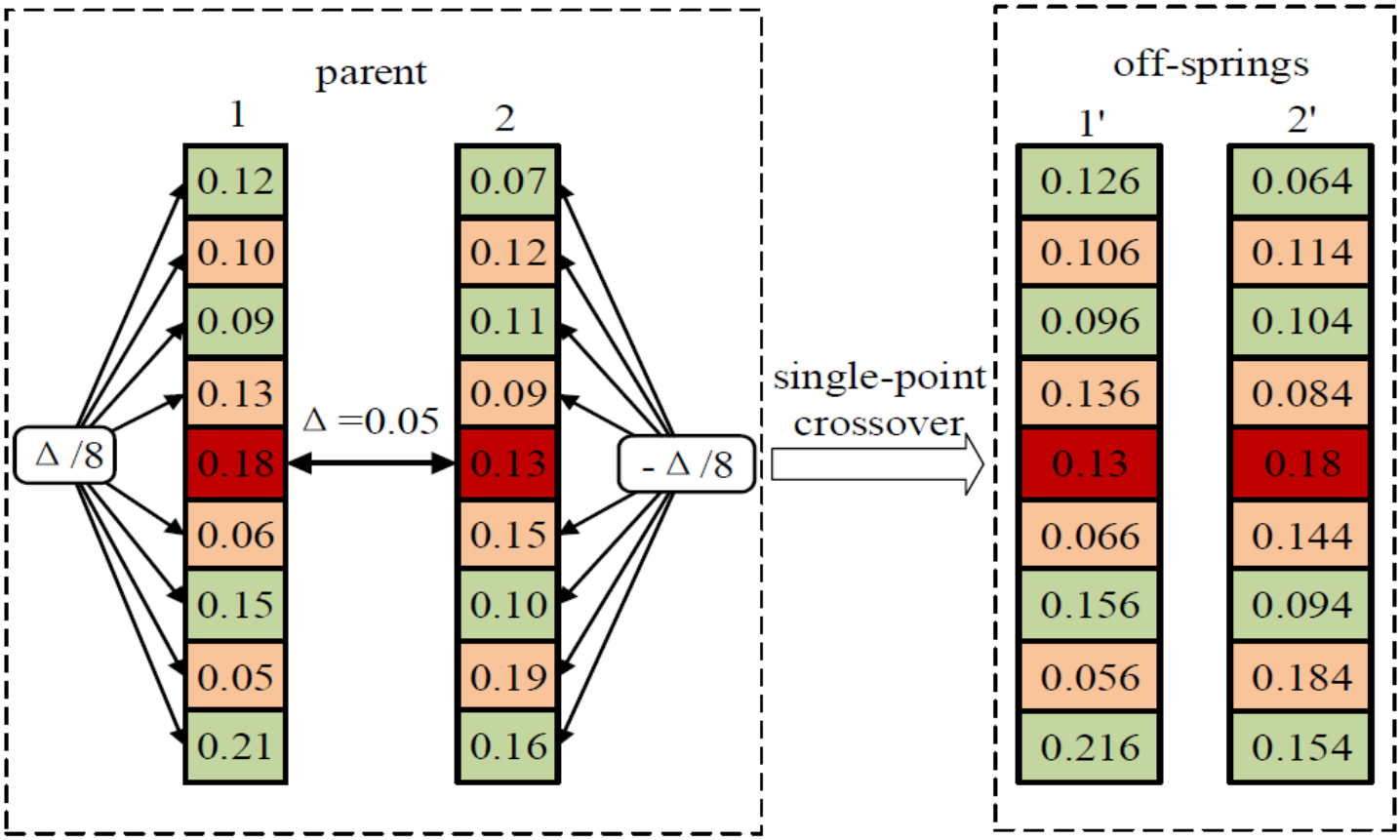

Figure 4

An example of single point crossover.

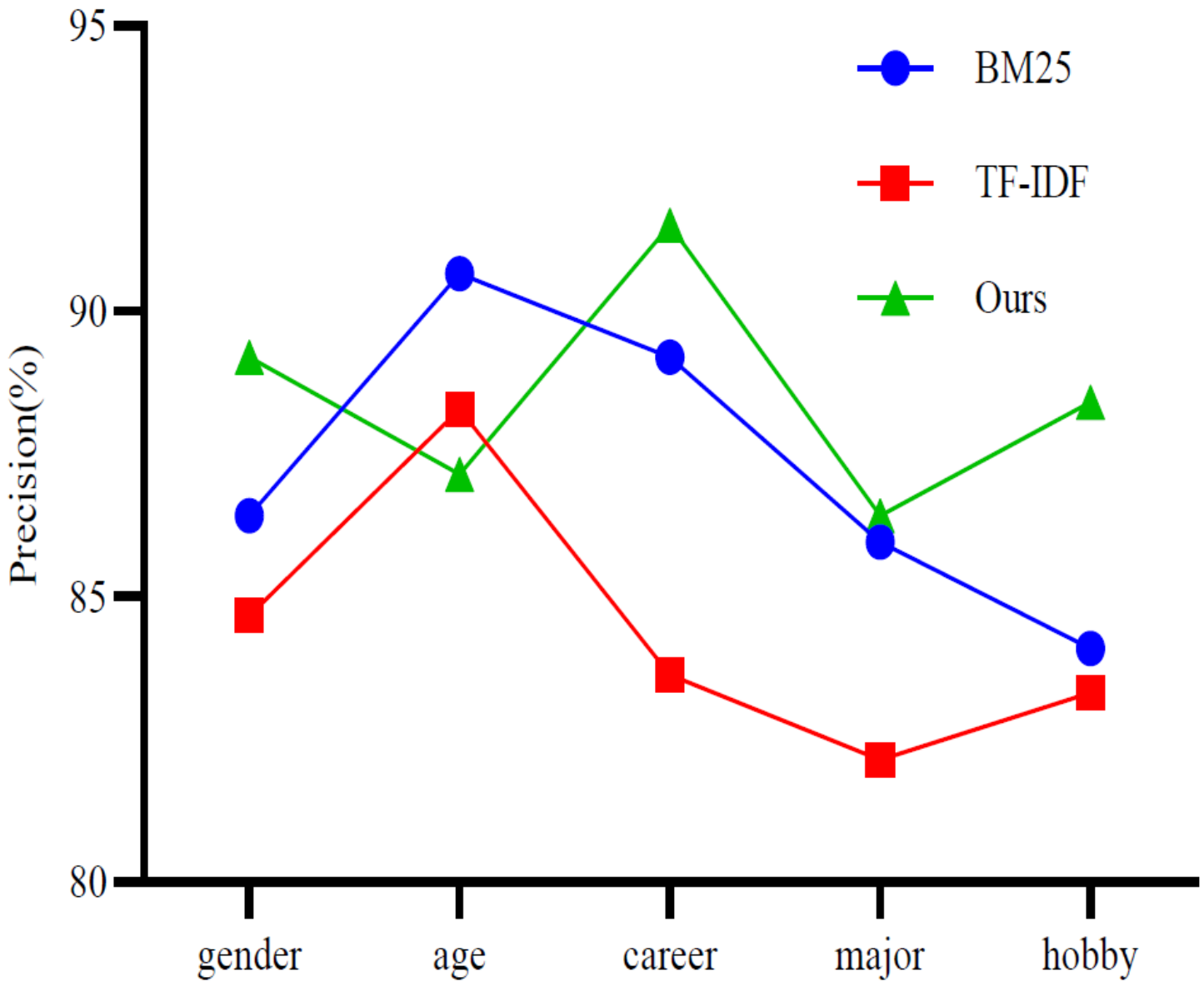

Figure 5

Accuracy comparison on dataset from Zhihu.

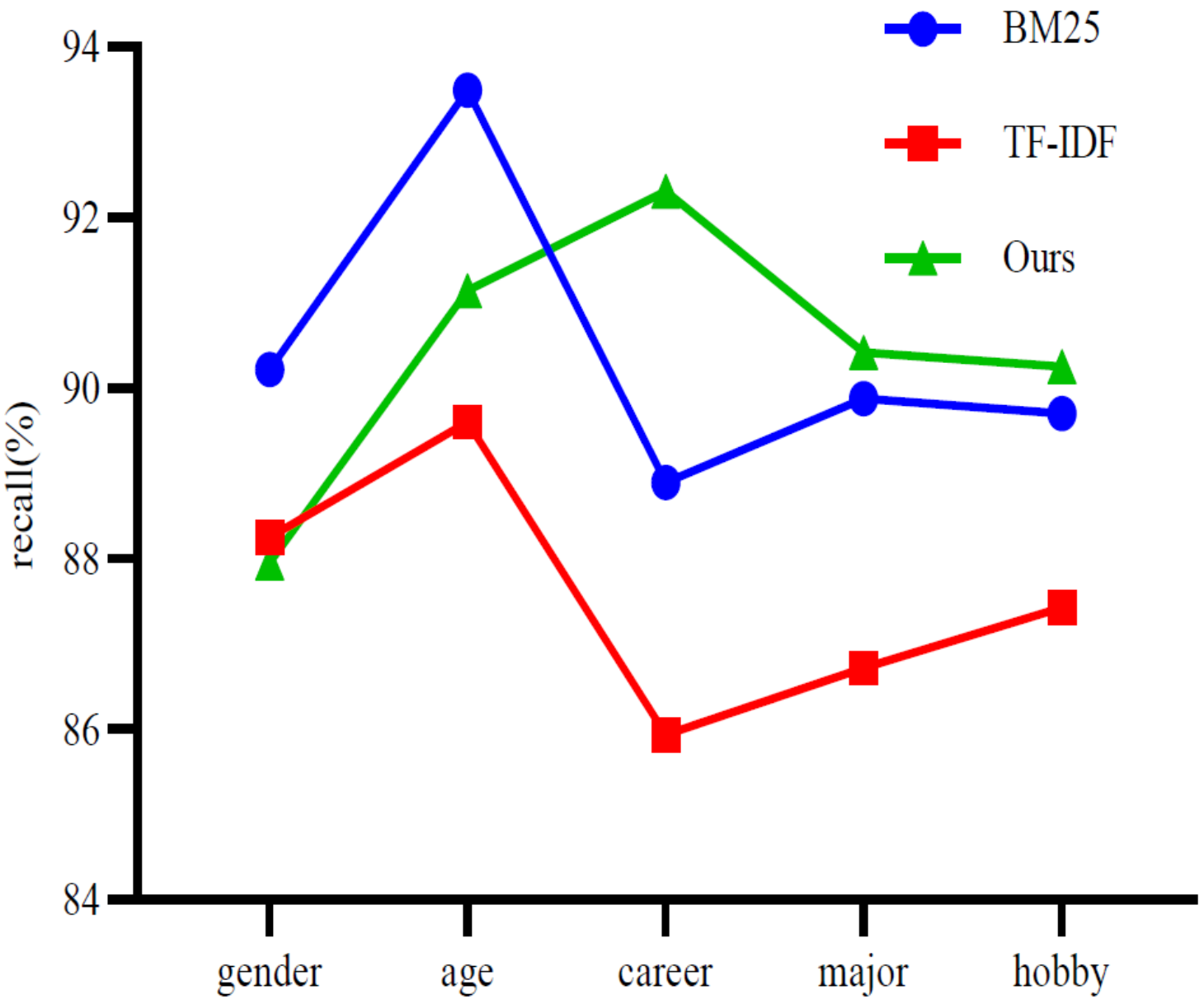

Figure 6

Recall comparison on dataset from Zhihu.

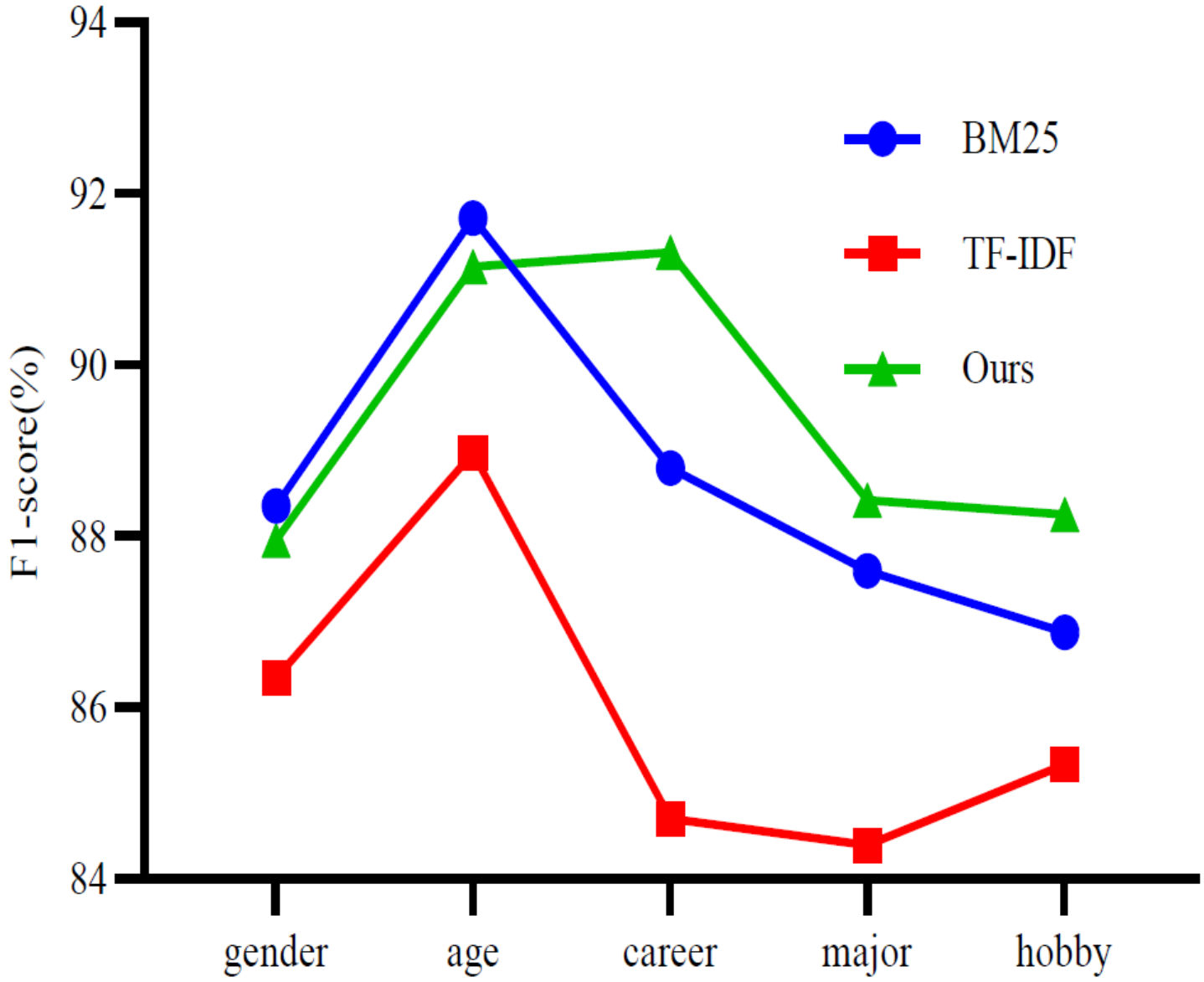

Figure 7

F1-score comparison on dataset from Zhihu.

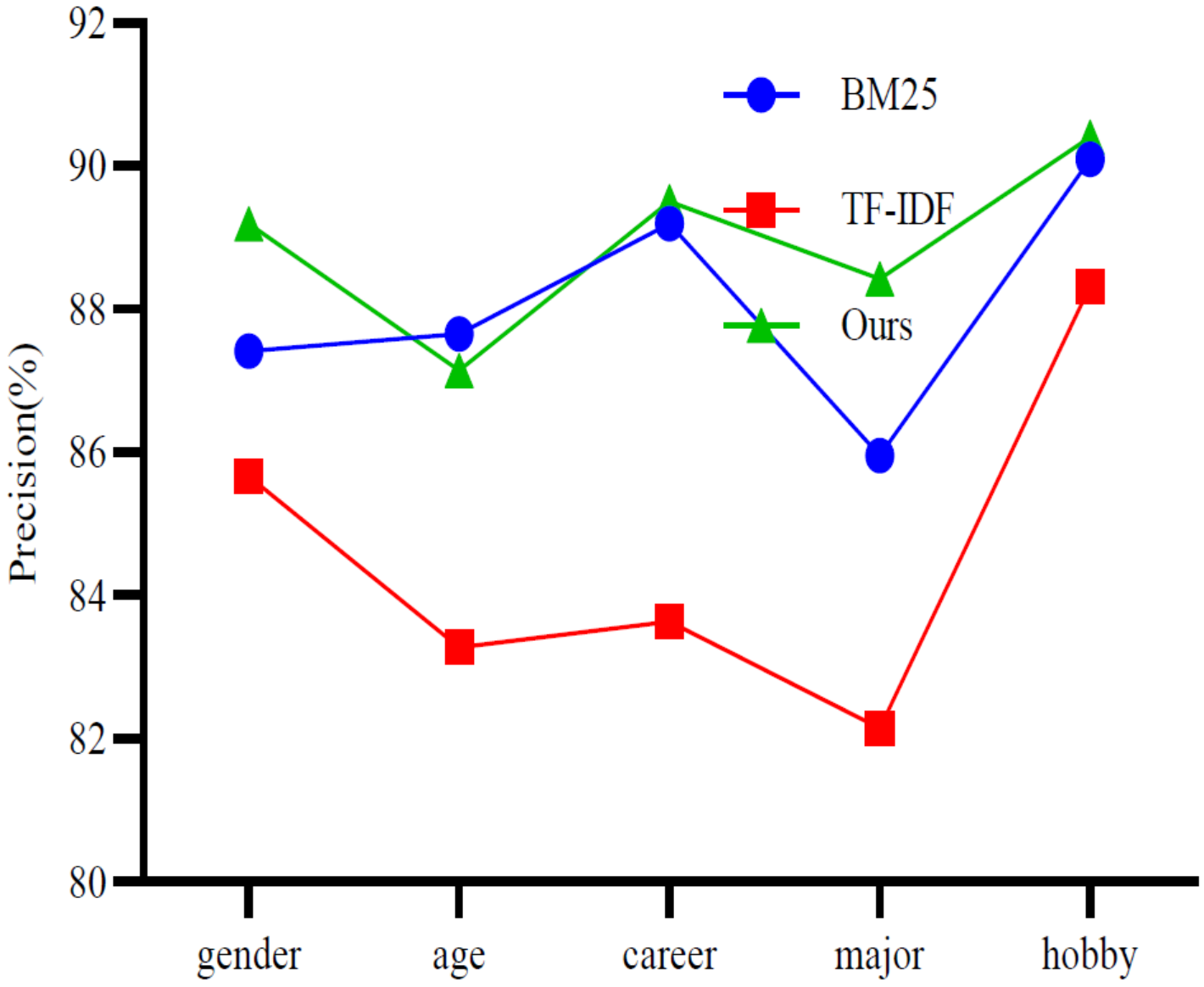

Figure 8

Accuracy comparison on dataset from Weibo.

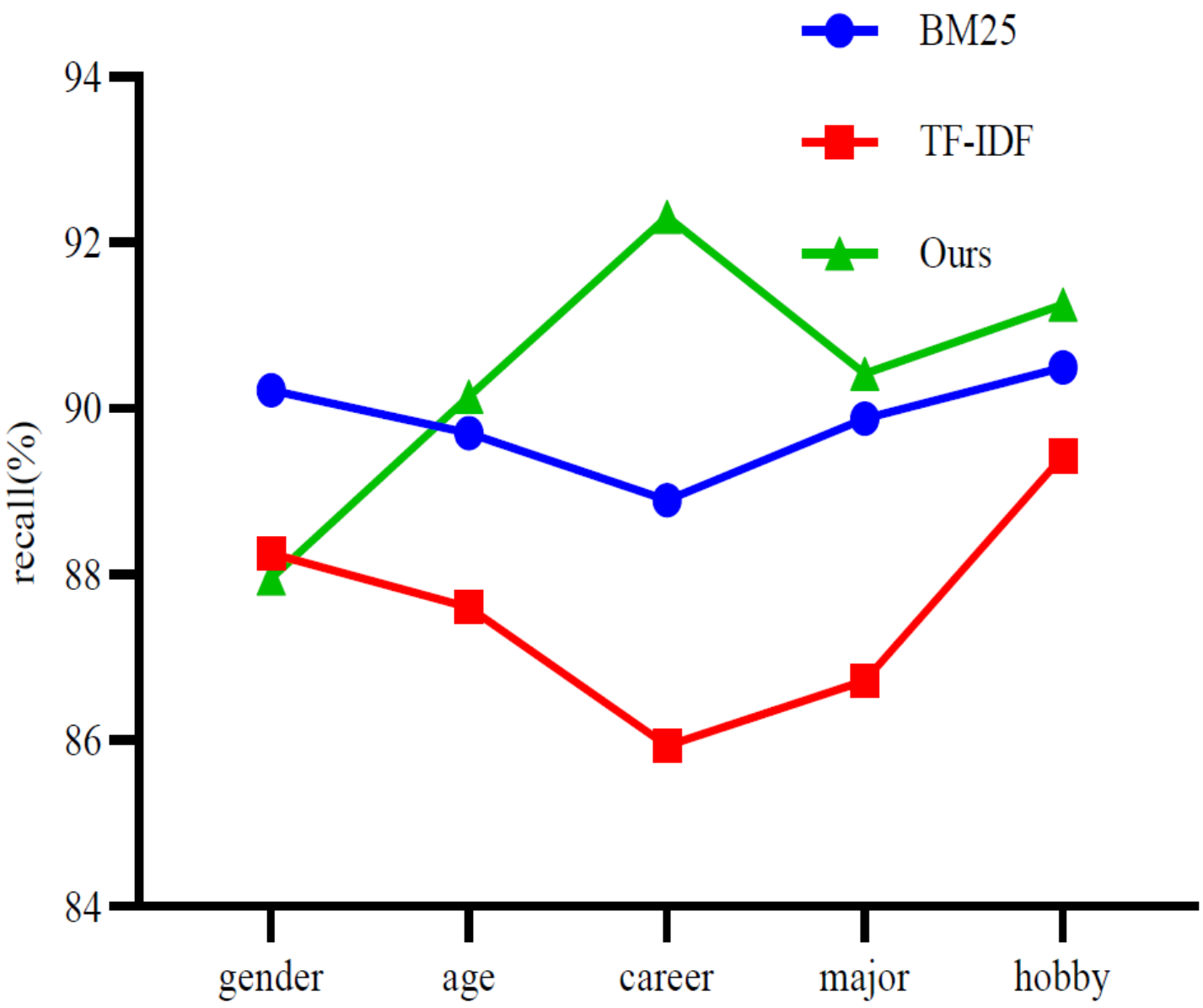

Figure 9

Recall comparison on dataset from Weibo.

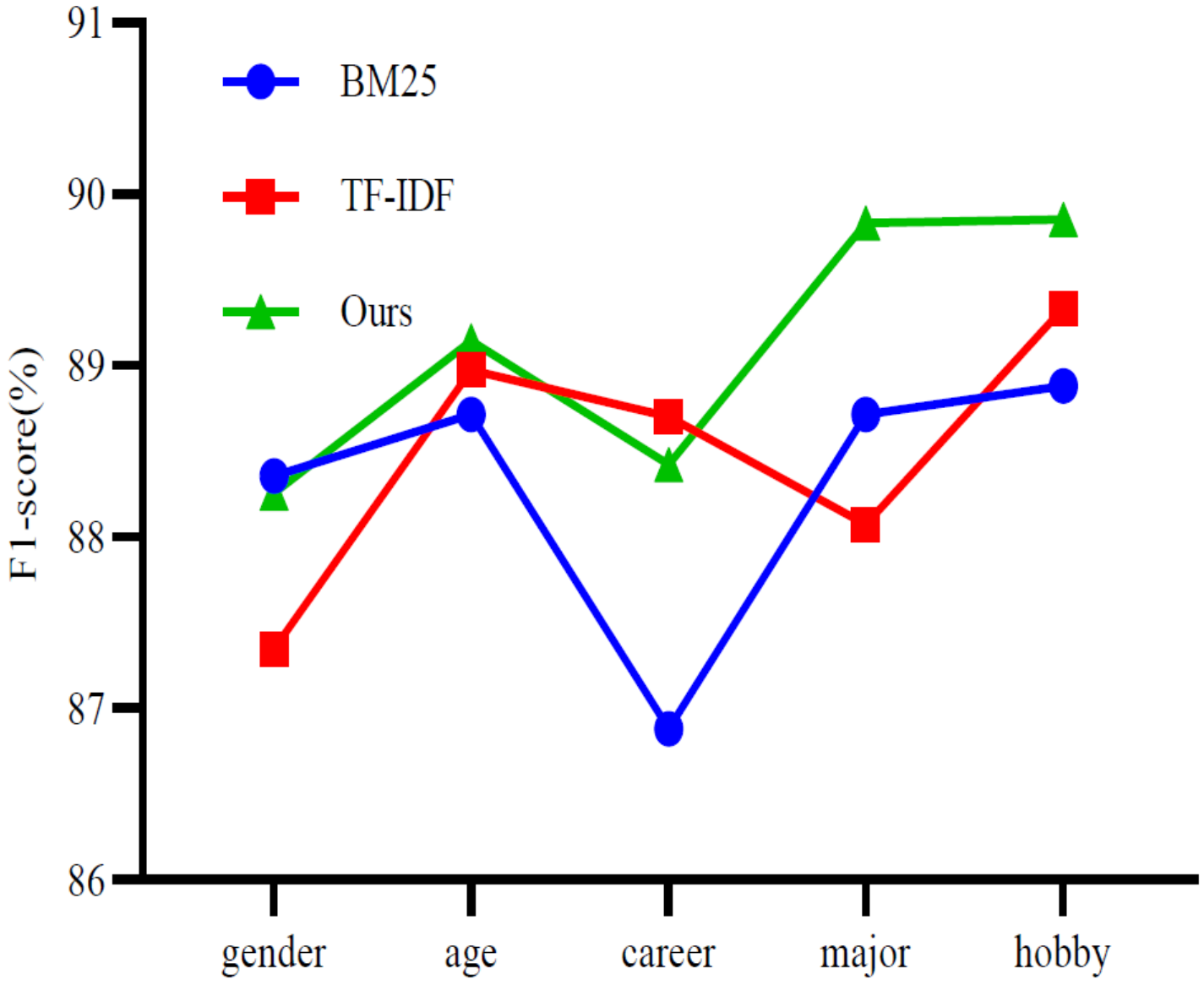

Figure 10

F1-score comparison on dataset from Weibo.