Personalized Search Using User Preferences on Social Media

Kyoungsoo Bok [1], **Jinwoo Song** [2], **Jongtae Lim** [2] and **Jaesoo Yoo** [2,*]

1 Department of Artificial Intelligence Convergence, Wonkwang University, Iksandae 460, Iksan 54538, Jeonbuk, Korea
2 Department of Information and Communication Engineering, Chungbuk National University, Chung-dae-ro 1, Seowon-Gu, Cheongju 28644, Chungbuk, Korea
* Correspondence: yjs@chungbuk.ac.kr; Tel.: +82-43-261-3230

Abstract: In contrast to traditional web search, personalized search provides search results that take into account the user's preferences. However, the existing personalized search methods have limitations in providing appropriate search results for the individual's preferences, because they do not consider the user's recent preferences or the preferences of other users. In this paper, we propose a new search method considering the user's recent preferences and similar users' preferences on social media analysis. Since the user expresses personal opinions on social media, it is possible to grasp the user preferences when analyzing the records of social media activities. The proposed method collects user social activity records and determines keywords of interest using TF-IDF. Since user preferences change continuously over time, we assign time weights to keywords of interest, giving many high values to state-of-the-art user preferences. We identify users with similar preferences to extend the search results to be provided to users because considering only user preferences in personalized searches can provide narrow search results. The proposed method provides personalized search results considering social characteristics by applying a ranking algorithm that considers similar user preferences as well as user preferences. It is shown through various performance evaluations that the proposed personalized search method outperforms the existing methods.

1. Introduction

With the advancement of the Internet and information systems, various data have been generated and shared on the web, and information search has become increasingly important [1,2]. When entering search terms into a web search engine, users hope that the results that they want are shown at the top of the list, allowing them to quickly find the desired web pages [3–5]. However, the traditional web search engine determines the importance of each web document that contains the search term through ranking algorithms to provide search results to users [6–8]. Therefore, most web search results do not provide customized search results because they determine common results to all users [9–12]. This problem arises because the method considers only the query form for the content-based document search without accurately understanding the search intentions of web users. That is, the traditional web searches do not provide search results suitable for the user's major preferences because they do not sufficiently reflect the user's query intention [13–16].

When the term 'Liverpool' is entered for a web search, results for 'Liverpool Football Club' and 'Liverpool City' are provided. In general, 'Liverpool Football Club' is at the top of the list, and a large number of related webpages are displayed. Such results are provided because the search engine seeks to show the results that users look for by considering the ambiguity and lexical ambiguity of the word. Because users often click on web pages associated with 'Liverpool Football Club', there are many web pages shown for it. Thus, the need for personalized search that provides results according to the individual users' preferences is increasing [17–20]. The personalized search can present the documents that

the user wants at the top of the search results according to the user's preference [21–23]. The traditional web search methods do not show the results that the user wants at the top, and additional user effort is needed for obtaining the desired information or avoiding unwanted information in many web search results. To solve this problem, a search method that provides documents containing the information that the user wants and places them at the top of the results is required. To provide a personalized search method that satisfies users, it is necessary to identify the exact meaning of the query that the user enters and determine what information the user needs according to his/her preference [24,25]. Therefore, for a personalized search, it is necessary to determine the meaning of the query using the user's preference.

To provide personalized search results, it is important to identify the preferences of the users. With the recent advancement of Internet technology and mobile devices, communication among users has become more active, and social media services have evolved as forums of communication. Internet technology allows users to access social media services quickly through mobile devices or the web and creates and accesses information quickly and conveniently [26,27]. The advancement of mobile devices has allowed users to access services anytime and anywhere. Social media services have been actively developed as a means of producing, consuming, and sharing information, and the number of users employing these services has increased rapidly [28–34]. Social media services involve two-way communication, in which a user becomes an information provider and a consumer simultaneously, whereas traditional media, such as newspapers, magazines, television, and radio, are one-way media, in which information producers deliver information to information consumers [35–38]. Social media are characterized by the rapid spread of information because users can produce, process, and share information themselves, and the processes are simple and convenient. Because of these characteristics, social media services have many users [39–41]. The amount of social data is rapidly increasing because of the increase in social media activity, and these data can be used to obtain user preferences that are useful for personalized search [42–44]. A user's social media activities can be analyzed to determine his/her fields of interest, and this information can be used to place the search results that reflect the user's preference at the top [45–50].

Studies have been performed on personalized search methods considering the preferences of social media users [51–54]. In [55], a method was proposed in which a query is accepted after identifying the user's preferences and a personalized search is executed using the social media analysis strategy. In [56], a method of classifying users' preferences by time was proposed, along with a personalized search method considering the network characteristics of social media. In previous studies, user preferences and social media network characteristics were employed to provide information that can satisfy users. However, except for [56], in which user preferences were classified by time, the user preferences were not updated. Furthermore, the method of [56] does not use time-based weights. Thus, different weights are not assigned for recent preferences and past preferences. Therefore, it is difficult to calculate the recent preferences of social media users accurately in the existing personalized search methods.

In this paper, we propose a personalized search method considering the user's recent preferences and similar users' preferences in social media environments. The recent preferences of users are considered to expose the information that the user wants at the top of the search results. The user's activities on social media are analyzed to determine his/her recent preferences, to which time-based weights are assigned for emphasizing the recent preferences. Furthermore, the preferences of similar users on social media are taken into account to consider new information, unknown information, and the case where the user information is unclear. Search results are provided considering the recent preferences of the user and the preferences of other users with high degrees of similarity, including users who are professionally active in the field where the user has recently shown interest and users who are linked as friends on social media. We propose a ranking method in which the user's recent preferences obtained from social media activities and the preferences of

similar users are employed to provide results. Our objective is to develop a personalized search method that improves the satisfaction and accuracy of the search results by updating user preferences via the determination of the user's recent preferences and considering the preferences of similar users.

The remainder of this paper is organized as follows. Section 2 introduces related works and presents the research problems. Section 3 describes the characteristics of the proposed personalized search method. Section 4 presents the experimental validation of the proposed method. Finally, Section 5 presents the conclusions.

2. Related Works

Studies have been actively conducted to increase the efficiency of web search, including personalized search. In particular, studies have been conducted on methods that use search record analysis to identify the user's preferences and then reflect it in search results to provide the user with desired information efficiently. Moreover, studies have been actively performed on search methods that use preferences of related users, such as similar users, users who engage in professional activities in a particular field, and users linked as friends. In [57], a method for finding experts on queries was proposed to obtain better answers compared with the traditional method of obtaining information. When a user sends a query, the proposed social query/answer system analyzes the user's social media and determines the user's ranking to deliver questions. However, it is difficult to identify the user's recent preferences, because the results for the current user's query are provided by aggregating all the activities of the user. In [58], a search algorithm was proposed that reflects the answerer's recent preferences to search for potential answerers who can best answer the question. It uses the bookmarking information provided by del.icio.us to obtain the user's recent preferences and considers the similarity between the user's query and preference to provide search results. However, when the degree of similarity between the user's query and the social annotation is low, appropriate search results cannot be provided. Furthermore, the proposed method cannot be applied to regular webpages, because the categories of the documents are unknown.

Recently, studies have been conducted on search methods employing social media, in which users exchange opinions or views with others. In [55], SonetRank was proposed, which provides web search results based on the preferences and feedback of users in a similar group. SonetRank accepts the user preference information in the form of a profile in advance as an input and then forms a group of similar users based on this information. The Social-Aware Search (SAS) was proposed, which analyzes the keywords searched and webpages viewed by the users belonging to the formed group to reflect the information related to the trends and fields of interest of the users in the group in the search results. It analyzes the preference for the documents that the user is typically interested in and places the user in a group of users with similar preferences. Then, when the user performs a search, the document viewing trend of the users in the group is given as a weight. In [59], a method was proposed in which fields of interest are identified according to the preferences of social media users and the user preferences are matched to the web search results to provide search results. After the user's preferences are identified in social media, the query is accepted, and a personalized search is executed according to the social media analysis strategy. However, most users' preferences change over time, and accordingly, their search preferences also change. Furthermore, traditional search methods provide limited results because they do not consider the preferences of similar users. In [60], a recommender system framework comprising a robust set of techniques was designed to provide mobile application developers with a specific platform. This framework consists of domain knowledge inference, profiling and preferencing, query expansion, and recommendation and information filtration to recommend and retrieve code snippets, Q&A threads, tutorials, libraries, and other external data sources and artifacts to assist developers with their mobile application development. The domain knowledge inference provides various semantic web technologies and lightweight ontologies. The

profiling and preferencing generated a new proposed time-aware multidimensional user modelling. The query expansion enhances the retrieved results by semantically augmenting users' query.

Most users' preferences change over time, and accordingly, so do their search preferences. Recently, studies have been conducted on search methods that use social media to identify and reflect the preferences of users that change over time. In [56], the use of profiles to classify the recent preferences of users by time was proposed, along with a personalized search method that considers the network characteristics of social media. When a user submits a query, results reflecting the user's preferences are provided. The user's preferences are built through click logs of the search results. Although a profile is employed in which the field of interest changes over time, the method has the following problems: there are no time-based weights for past and recent time periods, there are no exact time periods, and the user's preferences generated in real time on social media are not reflected. In [61], a search method was proposed that enhances reliability through an implicit information collection method using Skyline and receives feedback on the preference information about places to reflect it in searches. This method provides search results suitable for a moving user by including the time information in the query. It collects and analyzes social media postings, including a variety of location information, and the user requests search results via a search query to the server in a mobile environment. The server that receives the request analyzes the query, generates a candidate group, and assigns scores, and according to these scores, it provides search results to the user. Furthermore, it extracts the essential keywords, the time, and the user's current location from the user's query. It generates appropriate candidates, calculates popularity scores, and assigns weights based on the user's preference information. A Treatment Effect Pattern (TEP) was proposed to determine whether to take a treatment for personalized decision making [62]. TEP uses the local causal structure for unbiased Conditional Average Causal Effect (CATE) estimation in our problem setting. TEP uses a bottom-up search approach to represent treatment effect heterogeneity in data. Because the subgroup of the TEP is small, the subgroups are merged with other subgroups to make the TEP of the merged subgroup significant. A generalized TEP created by the merge process represents the two or more most specific TEPs. The discovery process minimizes heterogeneity within each subgroup represented by a pattern. The most specific pattern matching a person's situation is used for personalized decision making.

A study on event and topic detection has been proposed to provide context-aware services on social media. In [63], a public psychological pressure index was proposed to measure public opinion in social networks. The public psychological pressure index represents the status of public psychological pressure in relation to specific social events or topics to measure public psychological pressure in social networks. This index considers the probability distribution under the maximum entropy constraint condition. The public psychological pressure entropy is used as an important assessment quantitative indicator for public opinion analysis. Online learning comment is not only a textual evaluation, but also reflects changes in the learner's behavior, knowledge, and emotion. In [64], Social–Emotional Semantic Model (SESM) was proposed to extract a comment's social and emotional semantic meaning. SESM considers online learners' comments as semantics-based interaction. Keywords are extracted from the complex comments by the term frequency–inverse document frequency (TF-IDF). There are three relation types such as the user–evaluate–course relation, the user–reply–comment relation, and the user–post–comment relation. Double time series emotional analysis, which considers user-based time series algorithm and topic-based time series algorithm, analyzes and detects the emotional change within the current topic and all users. A peak-detection approach using social geo-tagged data was proposed to detect local events that occur in a specific region during a given time window [65]. This method detects events through space-time feature extraction and peak detection. Each space-time feature is then modeled as a time series. It extracts text features by considering only hashtags, which are a meaningful way to categorize messages. The entropy of the hashtags is extracted and is used as a feature to achieve

textual analysis. The baseline profile is computed through a scoring mechanism based on a statistics measure. The peak detection determines each element in the time series that is significantly distant from a baseline profile as a peak. The recommendation and information filtration provide personalized services to the designated users and answer users' queries with a minimum number of mismatches. In [66], an online clustering method was proposed to detect interesting topics in social data streams. To detect topics through incremental clustering, this method summarizes the investigated tweets into the cluster centroids. Each centroid is represented by multiple features, such as the timestamp of cluster generation and the timestamp of the last update, along with the set of terms appearing in the tweet. When a tweet arrives on a topic that has not appeared before, a new cluster is created. The online clustering incrementally groups similar tweets into the same cluster for the incoming social data streams. To assign a tweet to the centroid, the Jaccard similarity that takes into account both the cluster age and the terms occurring in the tweet is used.

AI-based approaches are being studied to provide optimized personalization services and accurate search results. In [67], AI search methods were analyzed for personalized cancer therapy synthesis to solve related problems occurring in clinical practice. A method to compute a safe and effective personalized treatment for Colorectal Cancer (CRC) was studied. A simulation-based, non-linear, constrained optimization problem was defined for automatically synthesizing personalized therapies. A word-distributed sensitive topic representation model, called WDS-LDA, was proposed for representing a topic in social networks [68]. The basic concept of WDS-LDA is that the distribution of words within a topic or among different topics has a great influence on the selection of topic expression words. WDS-LDA is based on the LDA model. The topic-word distribution acquired by the Latent Dirichlet Allocation (LDA) makes the representative words more important and makes the distinction among different topic words higher by considering the distribution of words between documents within a topic and among topics. To improve the precision of the subsequent topic detection and topic evolutionary analysis using the topic model, this method introduces the human cognitive ability and cognitive models to topic representation based on Hybrid human–AI (H-AI) and improves the precision of the subsequent topic detection and topic evolutionary analysis algorithms using the topic model. In [69], a framework for acquiring the domain of the textual content generated by users in an online social network was proposed. A Twitter mining incorporating machine learning is used for domain based classification of users and their textual content. This framework consists of three modules such as data collection and acquisition, features extraction, and machine learning. It constructs a data set by collecting the users' historical tweets containing public user content and metadata through the Twitter API. The data cleaning and integration techniques are applied to the collected dataset to ensure the certainty of the data. The features extraction extracts a list of user features. The new users' features are extracted in the user features extraction and the existing users' features are extracted in tweet features extraction. The machine learning module classifies users into political and non-political categories. To implement a computationally simple but effective approach, five classifiers are used: LR, SVM, top-down derivation-based decision tree (TD-DT), random forest-based decision tree (RF-DT), and gradient boosting-based decision tree (GB-DT).

3. Proposed Personalized Search Method

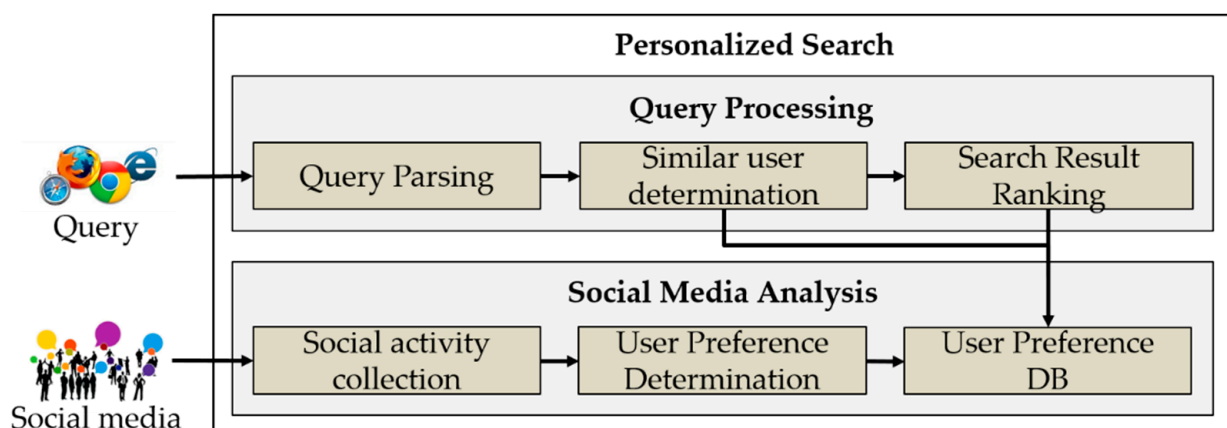### 3.1. Characteristics

With the advancement of information processing systems and various web services, countless pieces of information have been created, and the preferences of users are changing in a variety of ways. For identifying the fields in which the user is interested, it is important to determine the user's recent preferences. Because traditional personalized search methods do not consider the cases of information that is new to users, information unknown to users, and unclear information, it is necessary to develop a method that considers the user's recent preferences and similar users' preferences. In this paper, we propose a personalized search method that uses the social media user's recent preferences and similar users'

preferences. The proposed method categorizes keywords extracted using a text mining method from social media activity information of users and determines the preferences of users. Furthermore, it assigns time-based weights to the preferences of users, with smaller and larger weights assigned to older and newer preferences, respectively. It determines the preferences of similar users on social media, professional users of a particular field, and users linked as friends. Subsequently, it compares the extracted keywords, the user's preferences, and similar users' preferences to provide search results.

The proposed method consists of social media analysis and query processing, as shown in Figure 1. The user preference determination is a preprocessing operation in which social activities of social media users are continuously collected and analyzed. The collected social information is used to determine the user's preferences according to his/her social activities. Keywords are detected using a morpheme analyzer according to the user's social activities, and the user's preferences are determined from the detected keywords using TF-IDF. When the user submits a query, similar users are identified according to the query. For identifying similar users, the proposed method considers users who have similar preferences to the user according to the determined preferences of the user, users who are actively engaged in professional activities in a specific field according to the query, and users who are linked as friends to the user. Thus, it uses network characteristics of social media to consider users with strong social relationships. The ranking algorithm reflects the user's preferences in the web search results of the user query and then applies the preferences of similar users to present the search results. Furthermore, it uses the fractional cascading method to provide the search results in the descending order of the preferences matched between the user and similar users.
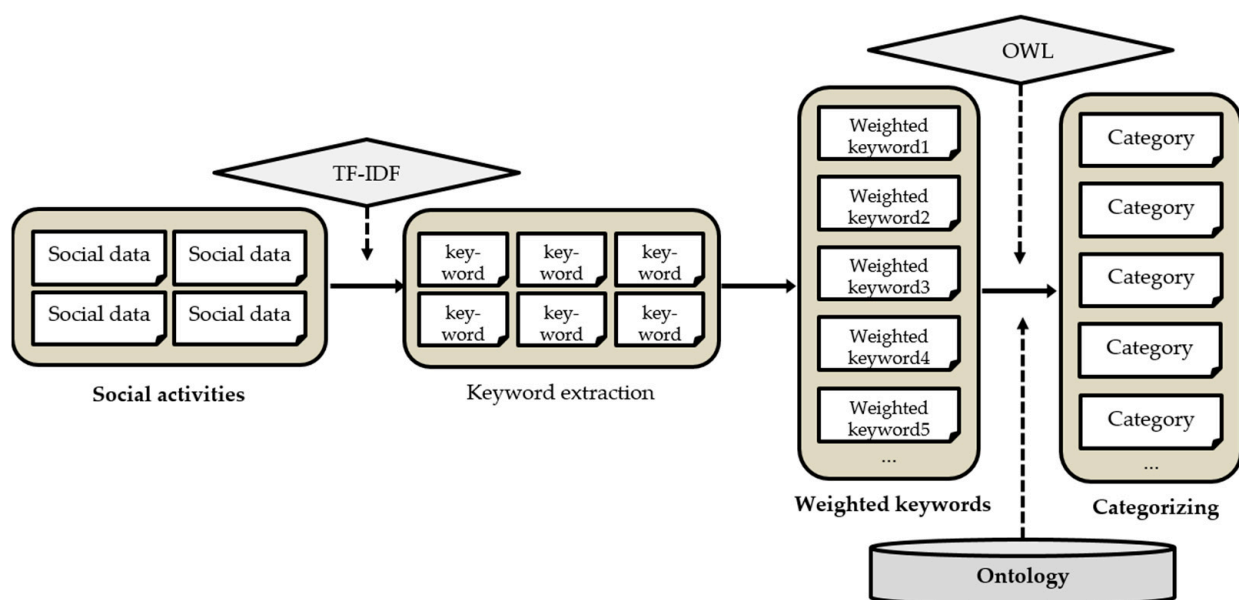

Figure 1. Overall architecture of the proposed search method.

3.2. User Preference Determination

The user preference determination is based on the postings that the user has created or shared on social media. The user's behavior of creating and sharing postings and writing comments can reflect his/her interests. Figure 2 shows the procedure of determining the user's preferences. The user preference determination is performed as a preprocessing task, and the activities of social media users are continuously collected and analyzed. The collected information is used to determine the user's preferences according to the user's social activities, and a morpheme analyzer is used to detect keywords from the user's social activity information. TF-IDF is a technique for determining the importance of words in documents in the field of information retrieval and text mining. Social media users express their opinions and thoughts in a variety of ways. Therefore, they include meaningless words in identifying user preferences. In other words, words generally used by users on social media or words that appear in all documents are not keywords that can grasp individual preferences. The proposed method applies TF-IDF to words extracted through a morpheme analyzer to extract keywords that reflect individual preferences

well among words contained in the user's social activity records. Keywords with high TF-IDF values are extracted as words that express the user's specific tendency well, and otherwise, they are recognized as meaningless words and removed. In TF-IDF, the term frequency (TF) indicates the frequency of a certain keyword appearing in the document and the inverse document frequency (IDF) indicates how many documents have a particular keyword among all the documents. For example, suppose that the user has engaged with 1000 documents, among which three documents contain the keyword 'Liverpool'. If the keyword 'Liverpool' appears three times in a document, the TF for this document is 3, and the IDF is 2.5215. Thus, the TF-IDF has a value of 7.5645. A higher TF-IDF score indicates that the keyword is more important. It is difficult to recognize the user's preferences according to simple keywords of interest. Therefore, a process of categorizing major keywords of interest obtained from social media activities is required. An ontology is constructed using the categories for classifying pages existing on social media [70], and the user's preferences are determined from the user's major keywords of interest using the constructed ontology and Web Ontology Language (OWL) [71].

Figure 2. Procedure of user preference determination.

It is difficult to identify the user's recent preferences if the user's profile is not updated periodically on social media. Therefore, the proposed method assigns time-based weights to the user's preferences to reduce the proportion of old preferences and increase the proportion of recent preferences. Accordingly, recent activities are analyzed to determine recent preferences. Identifying the user's preferences is an important element in analyzing the user's characteristics. The user's characteristics for determining the user's interests and preferences are crucial for providing a personalized search. Users have different weights for their fields of interest. Accordingly, different weights are assigned to the categorized preferences.

Equation (1) is used to calculate the scores for the different fields of interest. High frequencies of a certain keyword in the field of interest and the most recent search keyword indicate that the level of interest in that keyword is high. Therefore, time-based weights are assigned to the fields that the user is interested in, and an index is assigned to the keyword frequency to minimize the weakness of the mean. PW is a preference score that has categorized the extracted keyword. PW is represented by the frequency of the keyword of interest and the time-based weight. Here, n is the number of keywords for the field of

interest, tw_i is the time-based weight of the keyword, as given by Equation (2), and cnt_i is the frequency of the keyword.

$$PW = \frac{1}{n} \sum_{i=1}^{n} \left(tw_i \times e^{cnt_i} \right) \tag{1}$$

$$tw_i = -1.0506^t + 2 \tag{2}$$

3.3. User Similarity Determination

Applying personalized rankings has numerous risks. Even a user who usually prefers document 'A' may sometimes be interested in document 'B'. If only document 'A' is provided, the level of user satisfaction reduces. Similarly, although it is important to identify the user's preferences, if only the user's preferences are analyzed to perform a personalized search, the search results provide information that reflects the user's preferences from a narrow viewpoint [72]. The proposed method considers other users' preferences to provide meaningful documents, including content that users have not yet grasped or expressed in existing social activities. In order to consider other user preferences, similar users are discriminated in consideration of user preference similarity, expertise, and friend relationships. Collaborative filtering, which is frequently used in recommendation services, can provide information that cannot be provided by methods that consider only the user's personal preferences using similar users. Information of interest to users with similar user preferences may also be useful for a user who requests a recommendation service. Therefore, it is very important to identify users with similar preferences. There are experts among social media users who are interested in a specific field. We need to consider the user's expertise because experts can identify useful documents that the user needs. On social media, friends are users with similar or reliable preferences. We need to consider friend relationships because the friends can share or spread user's opinions and thoughts.

If users 'A' and 'B' are highly active in the sports, user 'A' can obtain information about the sports field from user 'B'. It is important to determine the user's preference for identifying similar users on social media. The user preference similarity between user u and user i is shown in Equation (3), where n represents the number of preference categories of the user, and k represents the number of preferences that match between the user and other users. A small score difference for a certain field indicates similar levels of interest in this field. Therefore, we calculate the difference in the preference score between the user and other users in a certain field and define the users with small differences as users with similar preferences.

$$IS_i = \frac{1}{n} \sum_{j=1}^{k} \left(1 - \left| PW_{uj} - PW_{ij} \right| \right) \tag{3}$$

We identify experts among social media users who can identify and provide useful documents. For example, if a certain user is highly engaged in professional activities for the sports, it can be assumed that he/she has considerable expertise in sports. If the preferences of users who have high levels of expertise are considered, useful information can be provided to the user. Equation (4) is used to calculate the expertise of users who are highly active for a specific preference, where n is the number of webpages that other users have engaged with, N_j is the number of participants for the webpage j, NR_{ij} is the number of recommendations, NS_{ij} is the number of times the webpage has been shared, and NC_{ij} is the number of comments. When the level of participation of other users in a document of a certain field is high, this indicates that the level of expertise of the user who created the document is high. Therefore, the recommendations, sharing, and comments are considered

according to the number of participants in the postings for the social media activities of users to determine their levels expertise and identify users with expertise.

$$PS_i = \frac{1}{n} \sum_{j=1}^{n} \frac{NR_{ij} + NS_{ij} + NC_{ij}}{N_j} \tag{4}$$

It is necessary to consider the preferences of users who are linked to the user as friends on social media. For example, if users 'A' and 'B' are linked as friends on social media, they may have a common preference. Therefore, if the preferences of close users are considered, useful information can be provided to the user. Equation (5) is used to calculate the scores of users who have strong social relationships with the user on social media, where j is the number of hops connected by a friendship. Because users close to the user can provide satisfactory information to the user, we apply six degrees of Kevin Bacon to assign larger weights to closer users on social media. If the social relationship exceeds seven hops, RS_i is set to 0.

$$RS_i = e^{1-\frac{1}{6}j} - 1 \tag{5}$$

User similarity is calculated by weighting IS_i, PS_i, and RS_i calculated in Equations (3)–(5). By assigning appropriate weights to the equations for determining the users who have similar preferences, users who are professionally active in a specific field, and users linked as friends, we can examine which element has a larger impact on the user. Here, $\alpha + \beta + \gamma = 0$. Users with high user similarity may influence determining necessary documents. Similar users' preferences are reflected when ranking search results.

$$US_i = \alpha IS_i + \beta PS_i + \gamma RS_i \tag{6}$$

Assuming that the preferences and social activities of users are shown in Table 1, let us determine the similarity of users A and B, and the similarity of users A and C. First, for the user similarity, we use the *TW*, which is the score for the category of interest to calculate *IS*. The *IS* score of users 'A' and 'B' is 0.5, and that of users A and C is 0.5. For experts of a certain category, we can use the social activity information to calculate *PS*. The *PS* scores of users B and C are 0.42 and 0.7, respectively. Suppose that the social relationships of users A, B, and C are 1 hop and that the values of the weights α, β, and γ are 0.45, 0.45, and 0.1, respectively. Then, the similarity score of users 'A' and 'B' and that of users 'A' and 'C' are 0.514 and 0.64, respectively. Therefore, user 'C' is the most similar user to user 'A'.

Table 1. Preferences and social activities of users.

User	TW	Keywords	Social Activities
A	Songs (0.5)	Acoustic collaboration, Sugarbowl, Park Won	Songs (likes: 50, comments: 50, shares: 5, participants: 500)
	Sports leagues (0.5)	Premier League, La Liga, Bundesliga	
	Movies (0.5)	The Great Gatsby, Les Misérables, Titanic	
B	Sports leagues (0.5)	Premier League, La Liga, Serie A	Sports leagues (likes: 100, comments: 100, shares: 10, participants: 500)
	Sports teams (0.5)	Liverpool, Barcelona, Chelsea	
	Sports stadiums (0.2)	Anfield, Camp Nou, Stamford Bridge	
C	Songs (0.5)	Acousweet, Lalasweet, Park Won	Songs (likes: 80, comments: 60, shares: 10, participants: 200)
	Music videos (0.5)	Russian Roulette, Cheer Up, TT	
	Music ranking (0.5)	Melon Chart, Gini Chart, Naver Chart	

The proposed method considers the preferences of the user and similar users to determine the fields of interest of other users in addition to the user's preferences. It can identify the user's preferences and similar users' preferences, whereby the user can get search results based on the preferences of other users in the field of interest. As discussed

previously in this section, it is important to determine the preferences of other users. In the personalized search, however, the user's preferences are more important than other users' preferences. The user's preferences should have larger weights than similar users' preferences. Therefore, the ranking algorithm prioritizes the user's preferences over other users' preferences.

3.4. Search Result Ranking

The ranking algorithm is the most important element in determining the quality of the search. Users desire more accurate information, and, therefore, the ranking is determined by combining the preferences of the user, similar users, professional users in a specific field, and users linked as friends to present social search results that reflect the user's interests. The proposed method employs the ranking algorithm to provide appropriate search results to the user according to the preferences of users collected continuously. The ranking algorithm involves a three-step process.

First, the user's search query is analyzed to determine the field of the query. For personalized search, the field of the query should be determined according to the preferences of users, and the field of the search query submitted by the user is determined to reflect the preferences of users who are similar or have relevant expertise. In the next step, the search results reflecting the user's preferences are sorted. For example, when a regular web search is performed using the keyword 'Liverpool', results for 'Liverpool Football Club' and 'Liverpool City' are provided. If the user has a stronger preference for the sports field, the search results for 'Liverpool Football Club' will be presented at the top. Lastly, search results reflecting the preferences of similar users are sorted once more. A user who typically prefers 'Liverpool Football Club' may sometimes have an interest in 'Liverpool City'. If search results are provided only for 'Liverpool Football Club', this may result in poor user satisfaction. Similarly, although it is important to identify the user's preferences, if only the user's preferences are analyzed to perform the personalized search, the search results will provide information that reflects the user's preferences from a narrow viewpoint. The search results that reflect the preferences of similar users are sorted to consider information that users have not found yet and the case where the user's information is unclear.

Algorithm 1 shows the ranking algorithm. First, keywords are extracted from the user's search query to determine the field of the search query. Then, web search results for the user query are fetched. Subsequently, keywords related to the user's preferences are matched via the fractional-cascading method [73]. A keyword related to the user's preferences is ranked higher if its number of matches is larger. Next, the keywords related to the preferences of similar users are matched via the fractional cascading method to determine the rankings once more, and the web search results are sorted again. The ranking algorithm considers the preferences of similar users after considering the user's preferences, and this is a constraint on the preferences of other users. When keywords related to the preferences of other users are matched in the ranking algorithm, they must be accompanied by keywords related to the user's preferences and then reflected in the search results. This is because the user's preferences should have larger weights than the preferences of similar users. If the keywords related to the preferences of similar users are not related to the user's preferences, results that are inappropriate for the user may be provided.

Algorithm 1: Fractional-cascading method-based ranking algorithm.

Q: user query
D: number of search results
U: set of users
US: user similarity
UK: set of user preference keywords
SK: set of preference keywords for similar users

---

counter [1. . . D] ← 0
determine a category related to Q based on U
determine similar users based on US
for each web search result R
for each user preference keyword UK
if (UK matches R)
counter[UK] ++
else
break
RR ← ranked search results
for each ranked search result RR
for each similar user preference keyword SK
if (SK matches RR)
counter[SK] ++
else
break
return personalized search results PR

---

4. Performance Evaluation

4.1. Evaluation Results

To prove the superiority of the proposed personalized search method, it was evaluated in comparison with a traditional web search method and the method proposed in [56]. Google search results were used for the traditional web search, and the method proposed in [56] is referred to as Topic-Driven SocialRank (TDSocialRank) for convenience. For performance evaluation, a total of 200 users were used by first selecting 30 users interested in soccer among Facebook users in Korea and then collecting additional users connected to friendship. We implemented the experiment using Java running on an Intel core i5-4460 CPU@3.20 GHz with 8 GB RAM. We collected posts, comments, and replies on Facebook for each user and extracted keywords through the Han Nanum Korean Morphological Analyzer [74]. We extracted an average of 10 keywords representing user preferences for each of the five categories that were mainly active for each user. In addition, we updated the user preference on a two-week basis and then gave a time-weight to the preference. To generate personalized search results, we applied a ranking algorithm to web search results that consider user preferences and similar user preferences. We evaluated precision, recall, F-measure, and G-measure by selecting specific users and comparing the results of searches visited by real users of the personalized search results.

Table 2 shows the search results obtained by applying the user's preferences and similar users' preferences in Table 3. We generated new search results using the results described above to measure how effectively the proposed method provided search results compared with the conventional search methods. In Table 2, the column on the left presents the Google search results, and the middle column presents the search results of TDSocialRank. The column on the right presents the search results of the proposed personalized search method. The proposed method displays the results reflecting the user's personal preferences higher on the results page compared with the conventional search methods.

Table 2. Search results for the Premier League.

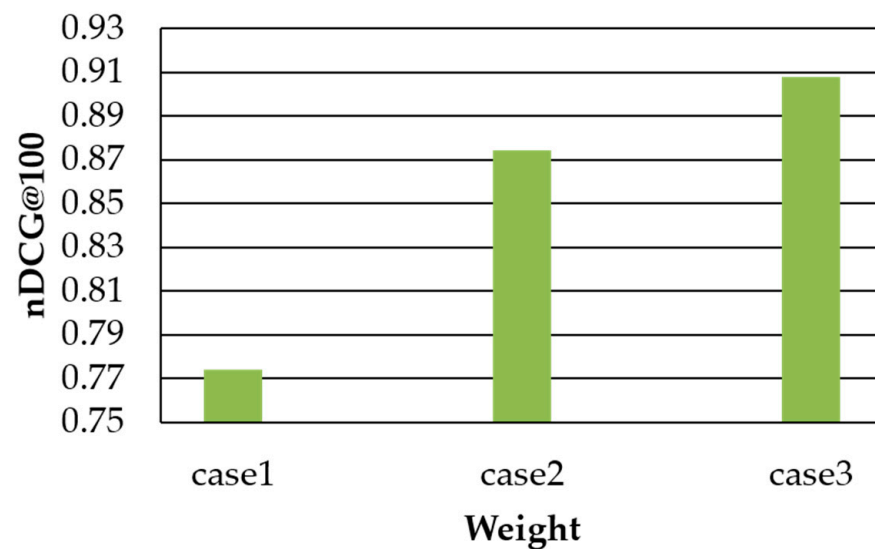| Web Search Results	TDSocialRank	Proposed Method
Premier League——Wikipedia, the free encyclopedia that anyone can edit	T.P.T.P-LIVERPOOL FC KOREAN FAN SITE	T.P.T.P-LIVERPOOL FC KOREAN FAN SITE
English Premier League——Namuwiki	[2015/16 season] Klopp, the manager who changes players best in the Premier League . . .	[2015/16 season] Klopp, the manager who changes players best in the Premier League . . .
Football: Premier League in 2015/2016 in real-time——-results, schedules, rankings . . .	11/12 English Premier League 2R——Arsenal vs Liverpool 110820 . . .	The Korea Times: [Duerden] Son Heung-min, why did he go to Tottenham?
English Premier League 2015–2016——data, all match schedules . . .	'Premier League', Tottenham, West Brom tie 1-1 . . . Son Heung-min in the second half	11/12 English Premier League 2R——Arsenal vs Liverpool 110820 . . .
Overseas football EPL match schedule/results | Daum Sports-Daum	[Premier League] Son Heung-min plays, Tottenham and West Brom tie . . .	'Premier League', Tottenham, West Brom tie 1-1 . . . Son Heung-min in the second half . . .
Premier League rankings——Goal.com	English Premier League——Namuwiki mirror	Sports & Entertainment: [Premier League] Son Heung-min "substitute player" . . . Tottenham vs. West Brom . . .
Why can't British people watch football games, not even once a month?: Europe . . .	[Premiers Skills English 14] Premier League 2014/15 season . . .	[English Premier League] Son Heung-min plays, Tottenham and West Brom tie . . .
English Premier League: SBS Sports	English Premier League tickets	Son Heung-min plays in the second half, Tottenham and WBA tie 1:1 . . . Far from the hope of winning the championship

Table 3. User's preferences and similar users' preferences in a specific field of interest.

Field of Interest	User	Similar User 1	Similar User 2
Sports teams	Liverpool (0.35) Klopp (0.35) Leicester (0.15) Son Heung-min (0.15)	Arsenal (0.35) Chelsea (0.35)	Tottenham (0.35)

To evaluate the proposed method, we compared the scores of similar user determination by changing the weights α, β, and γ in Equation (6). We used the nDCG for k shown recommendations (nDCG@k) to evaluate the satisfaction with each search result. nDCG is a metric that was developed because of the difficulty of providing differentiation based on rankings with the traditional accuracy and recall-based search engine evaluation method. In this evaluation method, the degree of satisfaction increases as the documents that the user wants are placed higher on the results page. Regarding the satisfaction score for each document in the search results, Equation (7) was used to calculate the satisfaction with the search result by assigning weights in the order of presenting the documents from top to bottom. In Equation (8), *IDCG* refers to the ideal score for the user's satisfaction with the search results. The *nDCG* score was calculated as the ratio of the *DCG* to the *IDCG*. Figure 3 shows the satisfaction evaluation results. In the proposed method, the weights were $\alpha = 0.33$, $\beta = 0.33$, and $\gamma = 0.33$ for case 1; $\alpha = 0.4$, $\beta = 0.4$, and $\gamma = 0.2$ for case 2; and $\alpha = 0.45$, $\beta = 0.45$, and $\gamma = 0.1$ for case 3. Users with similar preferences and users engaged in numerous professional activities had larger impacts on the search results than users with strong social relationships.

$$DCG_p = rel_1 + \sum_{i=2}^{p} \frac{rel_i}{log_2^i} \tag{7}$$

$$nDCG_p = \frac{DCG_p}{IDCG_p} \tag{8}$$

Figure 3. nDCG results based on weight changes.

The precision, recall, F-measure, and G-measure are important performance metrics in information searches. To confirm the superiority of the proposed method, we evaluated it in comparison with a traditional web search method and TDSocialRank [56] with regard to the precision, recall, F-measure, and G-measure. Here, the proposed method was experimentally evaluated by applying the weights of case 3, which exhibited the best performance among the cases, as shown in Figure 3. The precision is defined as the ratio of the number of relevant documents to the number of documents searched in the information search field, and the recall is defined as the ratio of the number of documents found in the search to the number of relevant documents in the information search field. The precision represented the number of documents related to the query in the search results, as indicated by Equation (9), and the recall represented the ratio of the number of relevant documents to the total number of documents as indicated by Equation (10). Here, true positive (TP) refers to the number of actually searching the documents that should be searched; false positive (FP) refers to the number of searching the documents that should not be searched; and false negative (FN) refers to the number of not searching among the documents that should have been searched. The F-measure is the harmonic mean of the precision and recall, as indicated by Equation (11), and the G-measure is the geometric mean of the precision and recall, as indicated by Equation (12).

$$Precison = \frac{TP}{TP + FP} \tag{9}$$
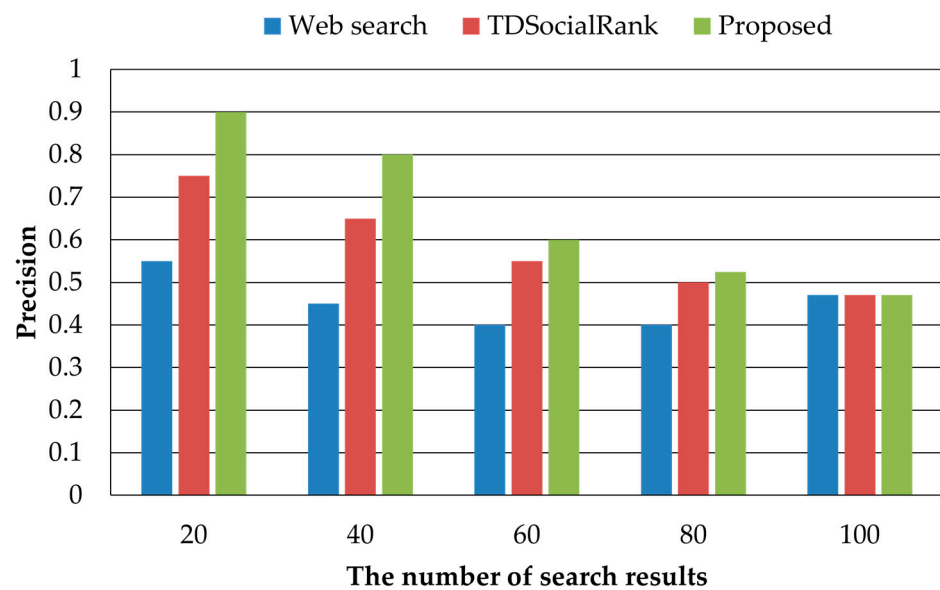
$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recal} \tag{11}$$

$$G - Measure = \sqrt{Precision \times Recall} \tag{12}$$

Figure 4 shows the results of comparing the precision among the proposed method and the existing methods. The typical web search exhibited the lowest precision because it did not consider the user's preferences. TDSocialRank considers the user's preferences, but it sometimes provides documents that should not be searched as search results, because recent user preferences are not properly reflected. The proposed method outperformed the existing methods because it provides results that properly reflect the user's recent preferences. As the number of search results increased, the precision decreased for all the methods. Nonetheless, the proposed method outperformed the existing methods. Compared with
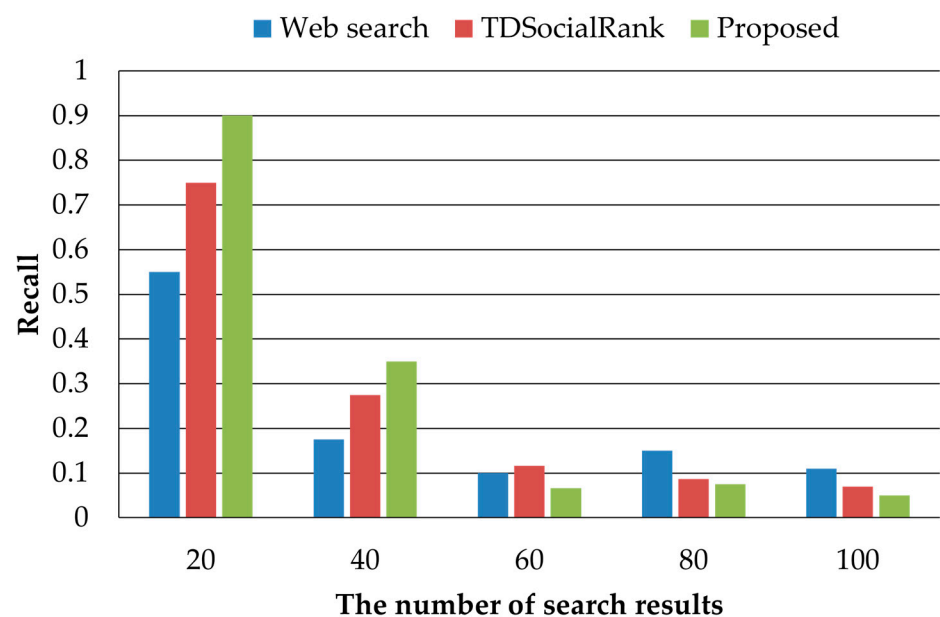
TDSocialRank, the proposed method exhibited a performance improvement of up to 23% when the number of search results was 40.

Figure 4. Precision.

Figure 5 shows the recall results for the proposed method and the existing methods. Similar to the case of precision, the regular web search method exhibited the worst recall performance. TDSocialRank exhibited better performance than the regular web search but was outperformed by the proposed method because it provided results that the user does not want. The proposed method provides only the search results that the user wants because it considers the user's recent preferences and the characteristics of users with similar preferences. Because the proposed method provides results considering the preferences of similar users, the performance declined when the number of search results increased. Nevertheless, it exhibited a performance improvement of up to 27% compared with TDSocialRank.

Figure 5. Recall.

Figure 6 shows the results of comparing the F-measure among the proposed method and the existing methods. The F-measure is a performance metric for combining the tradeoff relationship between the precision and recall. Even when the number of search results provided changed, the proposed method outperformed the regular web search with regard to the F-measure. However, for the proposed method, the recall decreased as the number of search results increased; thus, it was outperformed by TDSocialRank in some cases. When the number of the search results decreased, it exhibited a performance improvement of up to 26%.
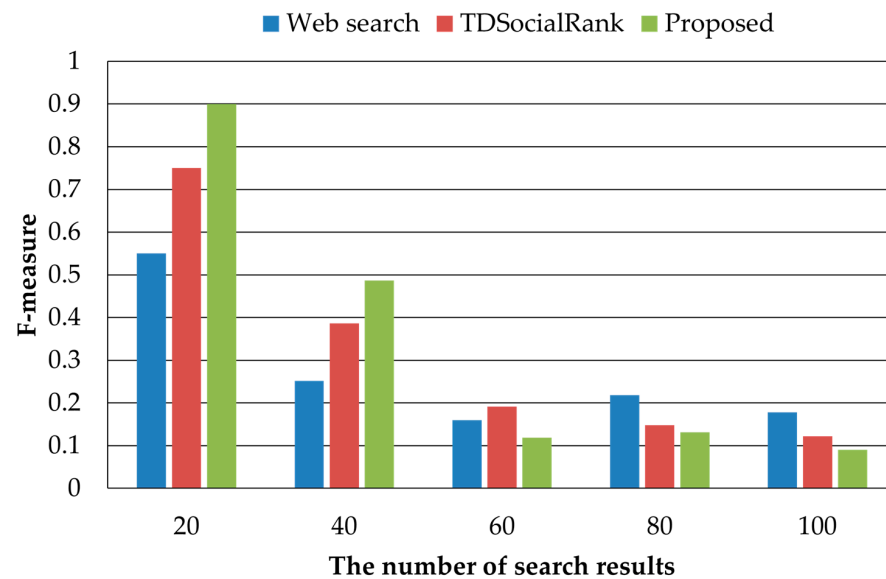

Figure 6. F-measure.

Figure 7 shows the results of comparing the G-measure between the proposed method and the existing methods. Even when the number of the search results changed, the proposed method outperformed the regular web search with regard to the G-measure. However, similar to the case of the F-measure, the proposed method had the problem that the recall decreased as the number of search results increased. Therefore, it was outperformed by TDSocialRank in some cases. When the number of the search results decreased, it outperformed TDSocialRank by up to 25%.
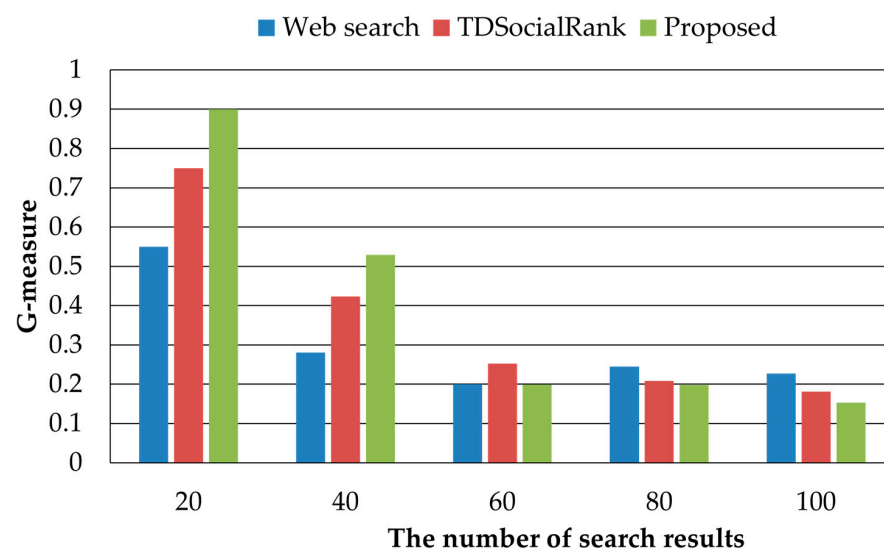

Figure 7. G-measure.

4.2. Discussion

General web search methods do not provide personalized search results because they provide search results based on the importance of popular documents. Personalized searches rank search results based on user preferences. The existing personalized search methods use search records and social media to determine personal preferences. In addition, search studies were conducted to identify users' preferences that change over time using social media and reflect them. Searches using preferences of users associated with users, such as similar users, users who engage in professional activities in a specific field, and users who are connected with friends, are being studied. However, the existing personalized search methods do not provide search results for content that users have not experienced because they identify personal preferences based on the user's past records. In the case of TDSSocialRank [56] used in performance comparison, the problem is that user preferences are determined through click logs, but they do not reflect user preferences generated in real time on social media. Furthermore, to provide search results based on personal preferences, search results cannot be extended considering other users' preferences.

When analyzing personal preferences by analyzing web usage records, only limited preferences can be determined. Since the user expresses personal opinions on social media, it is possible to grasp the user preferences when analyzing the records of social media activities. The proposed method collects user social activity records and determines keywords of interest using TF-IDF. In addition, time weights are given to the extracted keywords to reflect changes in individual preferences over time. Personalized search considering personal preferences does not provide search results with content that users have not experienced or grasped. To solve these problems, the proposed technique identifies similar users in consideration of user preferences, expertise, and friend relationships, and reflects similar users' preferences in search results. Accordingly, it is possible to provide a result reflecting not only the latest user preference but also the preference of other users.

Recently, in order to improve service satisfaction, various methods using context awareness or AI techniques are being conducted on social media. On social media, various opinions or thoughts are generated and shared among users. When analyzing such social media, events occurring around the user may be detected or the latest topics may be extracted. When using AI techniques, it is possible to grasp more improved preferences through machine learning. Furthermore, we can improve accuracy when applying AI techniques to event and topic discrimination. The proposed method analyzes social media through statistical analysis methods and provides personalized search results considering personal preferences and other user preferences. It is necessary to provide search results considering various characteristics as well as preference by applying a technique for determining events and topics. In addition, it is necessary to improve the performance of a personalized search method using the Graph Convolutional Network used as AI technology to determine personal preferences or applying techniques for context awareness.

5. Conclusions

In this paper, we proposed a personalized search method considering user preferences and similar user preferences. It is very important to consider user preferences for personalized searches. On social media, users can express their opinions and share or spread various opinions among users. In order to figure out the user preference considering the characteristics of social media, the user's activity records are collected on social media and the user's interest keywords are extracted using TF-IDF. Since user preferences change continuously over time, we assign time weights to keywords of interest, giving many high values to state-of-the-art user preferences. Providing personalized search results that only consider user preferences is useful for users, but it does not provide search results related to keywords that are not interested. The proposed method discriminates users with similar preferences and provides search results reflecting similar users' preferences. To demonstrate the superiority of the proposed method, we performed performance comparisons with a general web search method and TDSocialRank, a representative personalized

search method. In a performance evaluation, the proposed method achieved excellent performance with regard to precision, but in some cases, the recall decreased as the number of search results provided increased. Therefore, research is needed to improve the recall of personalized search. Recently, studies on event and topic determination techniques and studies to apply AI technology to achieve accuracy have been conducted for context awareness. In future research, we will conduct research to apply user preference and context identified through AI technology to personalized search. In addition, we will demonstrate the superiority of the proposed method through performance comparisons with recent works in various experimental environments.

Author Contributions: Conceptualization, K.B., J.S., J.L. and J.Y.; methodology, K.B., J.S. and J.L.; software, J.S.; validation, K.B., J.S. and J.Y.; formal analysis, K.B., J.S., J.L. and J.Y.; investigation, K.B., J.S. and J.L.; resources, J.S.; data curation, J.S.; writing—original draft preparation, K.B., J.S. and J.Y.; writing—review and editing, K.B., J.S., J.L. and J.Y.; visualization, J.S and J.L; supervision, J.Y.; project administration, K.B. and J.Y. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jansen, B.J.; Spink, A. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Inf. Process. Manag.* **2006**, *42*, 248–263. [CrossRef]
2. Yang, K. Information retrieval on the web. *Annu. Rev. Inf. Sci. Technol.* **2005**, *39*, 33–80. [CrossRef]
3. Ziakis, C.; Vlachopoulou, M.; Kyrkoudis, T.; Karagkiozidou, M. Important Factors for Improving Google Search Rank. *Future Internet* **2019**, *11*, 32. [CrossRef]
4. Schwartz, C. Web search engines. *J. Am. Soc. Inf. Sci.* **1998**, *49*, 973–982. [CrossRef]
5. Almukhtar, F.; Mahmoodd, N.; Kareem, S. Search engine optimization: A review. *Appl. Comput. Sci.* **2021**, *17*, 70–80.
6. Chung, F. A Brief Survey of PageRank Algorithms. *IEEE Trans. Netw. Sci. Eng.* **2014**, *1*, 38–42. [CrossRef]
7. Sharma, P.S.; Yadav, D.; Thakur, R.N. Web Page Ranking using Web Mining Techniques: A comprehensive survey. *Mob. Inf. Syst.* **2020**, *2022*, 7519573. [CrossRef]
8. Sankpal, L.J.; Patil, S.H. Rider-Rank Algorithm-Based Feature Extraction for Re-ranking the Webpages in the Search Engine. *Comput. J.* **2020**, *63*, 1479–1489. [CrossRef]
9. Agichtein, E.; Brill, E.; Dumais, S.T. Improving Web Search Ranking by Incorporating User Behavior Information. *SIGIR Forum* **2018**, *52*, 11–18. [CrossRef]
10. Plansangket, S.; Gan, J.Q. Re-ranking Google search returned web documents using document classification scores. *Artif. Intell. Res.* **2017**, *6*, 59–68. [CrossRef]
11. Liu, Y.; Wang, C.; Zhang, M.; Ma, S. User behavior modeling for better Web search ranking. *Front. Comput. Sci.* **2017**, *11*, 923–936. [CrossRef]
12. Jeh, G.; Widom, J. Scaling personalized web search. In Proceedings of the International World Wide Web Conference, Budapest, Hungary, 20–24 May 2003.
13. Singh, M.S.; Kumar, S. Personalized Web Search: A Survey. *Ind. Eng. J.* **2021**, *14*, 30–37.
14. Park, S.; Lee, W.; Choe, B.; Lee, S. A Survey on Personalized PageRank Computation Algorithms. *IEEE Access* **2019**, *7*, 163049–163062. [CrossRef]
15. Liu, J.; Liu, C.; Belkin, N.J. Personalization in text information retrieval: A survey. J. Assoc. *Inf. Sci. Technol.* **2020**, *71*, 349–369.
16. Benotman, H.; Maier, D. Comparing Personalized PageRank and Activation Spreading in Wikipedia Diagram-Based Search. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Champaign, IL, USA, 27–30 September 2021.
17. Liu, J.; Dou, Z.; Zhu, Q.; Wen, J. A Category-aware Multi-interest Model for Personalized Product Search. In Proceedings of the International World Wide Web Conferences, Virtual Event, Lyon, France, 25–29 April 2022.

18. Roul, R.K.; Sahoo, J.K. A novel approach for ranking web documents based on query-optimized personalized pagerank. *Int. J. Data Sci. Anal.* **2021**, *11*, 37–55. [CrossRef]

19. Nanda, A.; Omanwar, R.; Deshpande, B. Implicitly Learning a User Interest Profile for Personalization of Web Search Using Collaborative Filtering. In Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), Warsaw, Poland, 11–14 August 2014.

20. Wang, H.; He, X.; Chang, M.; Song, Y.; White, R.W.; Chu, W. Personalized ranking model adaptation for web search. In Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 July–1 August 2013.

21. Chebil, W.; Wedyan, M.O.; Lu, H.; Elshaweesh, O.G. Context-Aware Personalized Web Search Using Navigation History. *Int. J. Semant. Web Inf. Syst.* **2020**, *16*, 91–107. [CrossRef]

22. Hu, R.; Dou, W.; Liu, X.F.; Liu, J. Personalized Searching for Web Service Using User Interests. In Proceedings of the International Conference on Dependable, Autonomic and Secure Computing, Sydney, Australia, 12–14 December 2011.

23. Fathy, N.; Gharib, T.F.; Badr, N.L.; Mashat, A.S.; Abraham, A. A Personalized Approach for Re-ranking Search Results Using User Preferences. *J. Univers. Comput. Sci.* **2014**, *20*, 1232–1258.

24. Abri, S.; Abri, R.; Çetin, S. A Classification on Different Aspects of User Modelling in Personalized Web Search. In Proceedings of the International Conference on Natural Language Processing and Information Retrieval, Seoul, Korea, 18–20 December 2020.

25. Chawla, S. Web Page Recommender System using hybrid of Genetic Algorithm and Trust for Personalized Web Search. *J. Inf. Technol. Res.* **2018**, *11*, 110–127. [CrossRef]

26. Namisango, F.; Kang, K.; Beydoun, G. How the Structures Provided by Social Media Enable Collaborative Outcomes: A Study of Service Co-creation in Nonprofits. *Inf. Syst. Front.* **2022**, *24*, 517–535. [CrossRef]

27. Jain, P.K.; Pamula, R.; Yekun, E.A. A multi-label ensemble predicting model to service recommendation from social media contents. *J. Supercomput.* **2022**, *78*, 5203–5220. [CrossRef]

28. Asur, S.; Huberman, B. Predicting the future with social media. In Proceedings of the International Conference on Web Intelligence, Toronto, ON, Canada, 31 August–3 September 2010.

29. Mangold, W.G.; Faulds, D.J. Social media: The new hybrid element of the promotion mix. *Bus. Horiz.* **2009**, *52*, 357–365. [CrossRef]

30. Chow, W.S.; Chan, L.S. Social network, social trust and shared goals in organizational knowledge sharing. *Inf. Manag.* **2008**, *45*, 458–465. [CrossRef]

31. Qiu, L.; Sai, S.; Tian, X. TsFSIM: A three-step fast selection algorithm for influence maximisation in social network. *Connect. Sci.* **2021**, *33*, 854–869. [CrossRef]

32. Shioda, S.; Nakajima, K.; Minamikawa, M. Information Spread across Social Network Services with Non-Responsiveness of Individual Users. *Computer* **2020**, *9*, 65. [CrossRef]

33. Dang, V.T. Social networking site involvement and social life satisfaction: The moderating role of information sharing. *Internet Res.* **2021**, *31*, 80–99. [CrossRef]

34. Verduyn, P.; Gugushvili, N.; Massar, K.; Täht, K.; Kross, E. Social comparison on social networking sites. *Curr. Opin. Psychol.* **2020**, *36*, 32–37. [CrossRef] [PubMed]

35. Rousidis, D.; Koukaras, P.; Tjortjis, C. Social media prediction: A literature review. *Multim. Tools Appl.* **2020**, *79*, 6279–6311. [CrossRef]

36. Ali, K.; Hamilton, M.; Thevathayan, C.; Zhang, X. Social Information Services: A Service Oriented Analysis of Social Media. In Proceedings of the International Conference on Web Services, Seattle, WA, USA, 25–30 June 2018.

37. Zareie, A.; Sakellariou, R. Minimizing the spread of misinformation in online social networks: A survey. *J. Netw. Comput. Appl.* **2021**, *186*, 103094. [CrossRef]

38. Tian, Z.; Zhang, Z.; Xiao, D. Study on the Knowledge -Sharing Network of Innovation Teams using Social Network Analysis. In Proceedings of the International Conference on Enterprise Information Systems, Beijing, China, 8–11 June 2011.

39. Troudi, A.; Ghorbel, L.; Zayani, C.A.; Jamoussi, S.; Amous, I. MDER: Multi-Dimensional Event Recommendation in Social Media Context. *Comput. J.* **2021**, *64*, 369–382. [CrossRef]

40. Tareaf, R.B.; Berger, P.; Hennig, P.; Meinel, C. Cross-platform personality exploration system for online social networks: Facebook vs. Twitter. *Web Intell.* **2020**, *18*, 35–51. [CrossRef]

41. Ellison, N.B.; Vitak, J.; Gray, R.; Lampe, C. Cultivating Social Resources on Social Network Sites: Facebook Relationship Maintenance Behaviors and Their Role in Social Capital Processes. *J. Comput. Mediat. Commun.* **2014**, *19*, 855–870. [CrossRef]

42. Bok, K.; Noh, Y.; Lim, J.; Yoo, J. Hot topic prediction considering influence and expertise in social media. *Electron. Commer. Res.* **2021**, *21*, 671–687. [CrossRef]

43. Choi, D.; Park, S.; Ham, D.; Lim, H.; Bok, K.; Yoo, J. Local Event Detection Scheme by Analyzing Relevant Documents in Social Networks. *Appl. Sci.* **2021**, *11*, 577. [CrossRef]

44. Tabrizi, S.A.; Shakery, A.; Tavallaei, M.A.; Asadpour, M. Search Personalization Based on Social-Network-Based Interestedness Measures. *IEEE Access* **2019**, *7*, 119332–119349. [CrossRef]

45. Bok, K.; Song, H.; Choi, D.; Lim, J.; Park, D.; Yoo, J. Expert Recommendation for Answering Questions on Social Media. *Appl. Sci.* **2021**, *11*, 7681. [CrossRef]

46. Bok, K.; Ko, G.; Lim, J.; Yoo, J. Personalized content recommendation scheme based on trust in online social networks. *Concurr. Comput. Pract. Exp.* **2020**, *32*, e5572. [CrossRef]

47. Kurz, C.F.; König, A.N. Predicting time preference from social media behavior. *Future Gener. Comput. Syst.* **2022**, *130*, 155–163. [CrossRef]

48. Bok, K.; Lee, S.; Choi, D.; Lee, D.; Yoo, J. Recommending personalized events based on user preference analysis in event based social networks. *Electron. Commer. Res.* **2021**, *21*, 707–725. [CrossRef]

49. Dridi, A.; Slimani, Y. Leveraging social information for personalized search. *Soc. Netw. Anal. Min.* **2017**, *7*, 16. [CrossRef]

50. Xu, J.; Fu, X.; Wu, Y.; Luo, M.; Xu, M.; Zheng, N. Personalized top-n influential community search over large social networks. *World Wide Web* **2020**, *23*, 2153–2184. [CrossRef]

51. Bouadjenek, M.R.; Hacid, H.; Bouzeghoub, M. Sopra: A new social personalized ranking function for improving web search. In Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 July–1 August 2013.

52. Zhou, D.; Lawless, S.; Wu, X.; Zhao, W.; Liu, J. Enhanced Personalized Search using Social Data. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 1–4 November 2016.

53. Maalej, M.; Mtibaa, A.; Gargouri, F. Ontology-based User Model for Personalized Search in a Social Network. *Res. Comput. Sci.* **2018**, *147*, 87–106. [CrossRef]

54. Khalifi, H.; Dahir, S.; Qadi, A.E.; Ghanou, Y. Enhancing information retrieval performance by using social analysis. *Soc. Netw. Anal. Min.* **2020**, *10*, 24. [CrossRef]

55. Kashyap, A.; Amini, R.; Hristidis, V. SonetRank: Leveraging social networks to personalize search. In Proceedings of the International Conference on Information and Knowledge Management, Maui, HI, USA, 29 October–2 November 2012.

56. Kim, Y.A.; Park, G.W. Topic-Driven SocialRank: Personalized search result ranking by identifying similar, credible users in a social network. *Knowl. Based Syst.* **2013**, *54*, 230–242. [CrossRef]

57. Horowitz, D.; Kamvar, S.D. The anatomy of a large-scale social search engine. In Proceedings of the International Conference on World Wide Web, Raleigh, NC, USA, 26-30 April 2010.

58. Bao, S.; Xue, G.; Wu, X.; Yu, Y.; Fei, B.; Su, Z. Optimizing web search using social annotations. In Proceedings of the International Conference on World Wide Web, Banff, AB, Canada, 8–12 May 2007.

59. Shafiq, M.O.; Alhajj, R.; Rokne, J.G. On personalizing Web search using social network analysis. *Inf. Sci.* **2015**, *314*, 55–76. [CrossRef]

60. Abu-Salih, B.; Alsawalqah, H.; Elshqeirat, B.; Issa, T.; Wongthongtham, P.; Premi, K. Toward a Knowledge-based Personalised Recommender System for Mobile App Development. *J. Univers. Comput. Sci.* **2021**, *27*, 208–229. [CrossRef]

61. Bok, K.; Lim, J.; Ahn, M.; Yoo, J. A Social Search Scheme Considering User Preferences and Popularities in Mobile Environments. *KSII Trans. Internet Inf. Syst.* **2016**, *10*, 744–768.

62. Li, J.; Liu, L.; Zhang, S.; Ma, S.; Le, T.D.; Liu, J. Causal heterogeneity discovery by bottom-up pattern search for personalised decision making. *Appl. Intell.* **2022**, 1–15. [CrossRef]

63. Zhang, H.; Jin, R.; Zhang, Y.; Tian, Z. A Public Psychological Pressure Index for Social Networks. *IEEE Access* **2020**, *8*, 23457–23469. [CrossRef]

64. Weng, J.; Gan, W.; Ding, G.; Tian, Z.; Gao, Y.; Qiu, J. SESM: Emotional Social Semantic and Time Series Analysis of Learners' Comments. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Toronto, ON, Canada, 11–14 October 2020.

65. Comito, C.; Falcone, D.; Talia, D. A Peak Detection Method to Uncover Events from Social Media. In Proceedings of the IEEE International Conference on Data Science and Advanced Analytics, Tokyo, Japan, 19–21 October 2017.

66. Comito, C.; Pizzuti, C.; Procopio, N. Online Clustering for Topic Detection in Social Data Streams. In Proceedings of the IEEE International Conference on Tools with Artificial Intelligence, San Jose, CA, USA, 6–8 November 2016.

67. Esposito, M.; Picchiami, L. A Comparative Study of AI Search Methods for Personalised Cancer Therapy Synthesis in COPASI. In Proceedings of the International Conference of the Italian Association for Artificial Intelligence, Milan, Italy, 1–3 December 2021.

68. Han, W.; Tian, Z.; Zhu, C.; Huang, Z.; Jia, Y.; Guizani, M. A Topic Representation Model for Online Social Networks Based on Hybrid Human-Artificial Intelligence. *IEEE Trans. Comput. Soc. Syst.* **2021**, *8*, 191–200. [CrossRef]

69. Abu-Salih, B.; Wongthongtham, P.; Chan, K.Y. Twitter mining for ontology-based domain discovery incorporating machine learning. *J. Knowl. Manag.* **2018**, *22*, 949–981. [CrossRef]

70. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]

71. Bechhofer, S. OWL: Web Ontology Language. In *Encyclopedia of Database Systems*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2018.

72. Collins-Thompson, K.; Bennett, P.N.; White, R.W.; Chica, S.; Sontag, D.A. Personalizing web search results by reading level. In Proceedings of the ACM Conference on Information and Knowledge Management, Glasgow, UK, 24–28 October 2011.

73. Chazelle, B. Guibas, L.J. Fractional Cascading: I. A Data Structuring Technique. *Algorithmica* **1986**, *1*, 133–162. [CrossRef]

74. HanNanum Korean Morphological Analyzer. Available online: https://rdrr.io/github/haven-jeon/KoNLP/man/MorphAnalyzer.html (accessed on 21 July 2021).