Improving Online Dating Search Result Quality with Inverse Profile Matching and Activity Weighting

by Kenneth A. Granderson

Table of Contents

Abstract

There are hundreds, if not thousands, of Internet web sites set up to assist individuals in finding suitable partners for dating and/or marriage. While all of these sites are based on the concept of all potential matches completing an information profile containing a variety of relevant attributes, very few sites do any more than the most trivial information processing in order to assist users to find potential matches.

This paper describes a new technique for dramatically improving the quality of the matches returned from a dating service search process.

Terminology

In the interest of clarity and accurate understanding, this section defines several terms that will be used in this paper:

Profile

A Profile is a set of information that an online dating site user enters on the site that describes them in a number of ways. Profiles generally include their physical characteristics, interests and online dating objectives.

Attribute

An Attribute is an individual piece of information in a Profile, such as the person's height, weight, gender or location.

User

Generic term for anyone using the dating site, regardless of their role.

Seeker

The Seeker is the person initiating a search for a potential date.

Date

A Date (potential Date) is a another user of the dating site who has been identified as a potential Match for the Seeker.

Match (noun)

A Match is a person who meets the Seeker's dating objectives.

Match (verb)

Matching is the process of performing searches and making contact with another dating site user to determine if the parties are Matches for each other.

Search Population

The Search Population is the entire database of potential Matches in the online dating system.

Profiles

Dating profiles can be thought of as a collection of data attributes concerning the user who has created the profile. Attributes can take on several different forms that define their utility in the matching process (note: all attribute terminology is my own):

1. **Single Exact Attributes:**
   These are attributes that have discrete values and need to be matched exactly. An example is 'gender' (in sites whose objective is to match exactly 2 non-bisexual people). If I am seeking a woman, then the matching process should only include matches where the gender attribute is equal to female.
2. **Multiple Exact Attributes:**
   These are attributes that have discrete values, but one or more option may be acceptable. An example is 'religion.' In the predominantly Christian USA, the religion of Christianity is often broken down into several denominations in dating site profiles, and many seekers looking for Christians will be open to more than one denomination. The use of these kinds of attributes is determined by the seeker, as other seekers may prefer to date people who are practitioners of certain religions, they may prefer to avoid certain religions, or if they are part of an orthodox faith, they may treat religion as an Exact Attribute and only seek those of the same religion.
3. **Fuzzy Attributes:**
   These are attributes where a range of values are either supplied or are acceptable. An example is 'weight.' While user profiles may contain a discrete value or a range for user's weights, it is unlikely that seekers will be looking for a person of an exact weight. Rather, if weight is an issue, there will be a range of weights that the seeker will consider most desirable.
4. **Narrative Attributes:**
   These are attributes that contain narrative text entered by the profile creator. Unless the system controls data entry into these attributes, Narrative Attributes are generally only useful to the matching process if used in full-text searches.
5. **Media Attributes:**
   These are attributes that take the form of photographs, video or audio. While at least one major dating site provides a service that attempts to use pictures as searchable data, their implementation does not appear to be scalable, as it requires human intervention. Until / unless face recognition software becomes readily available, inexpensive and accurate even with amateur photographs taken in natural settings, media attributes are not likely to be viable for automated matching, and are thus not considered in this analysis.

As certain attributes will be more important than others in determining the quality of a match, it is possible to assign different weights to different attributes to produce more meaningful results. For example, the attribute 'religion' is probably more important to most users than the attribute 'eye color.' The optimal use of weighting is to enable the user to assign weights to attributes based on their unique preferences.

Conventional Matching

While none of the major dating sites publish their matching algorithms, anecdotal analysis of profile inputs and search results on a variety of sites leads to the following observations and conclusions:

1. Most sites offer only the most rudimentary matching on Single and Multiple Exact Attributes only.
2. Most sites implement Fuzzy Attributes as lists of Multiple Attributes rather than using true fuzzy search methodologies.
3. Only a few sites offer any user-defined attribute weighting capabilities.
4. Only a few sites compare the seeker's profile with the match result profiles and attempt to calculate match degrees of match relevance.

Conventional Matching's Biggest Problem

Regardless of the level of functionality offered in current matching sites, all of them (observed to date) suffer from a common, major flaw, which is that the matching process is performed from the perspective of the seeker, rather than the match. While this process can result in potential matches that are desirable to the seeker, as successful matching requires attraction from both parties, the quality of the match is equal or more dependent on the perspective of the match result (who is being sought) than the perspective of the seeker.

As an example, I think that Halle Berry is one of the most beautiful women alive, so I search for women who have her kinds of physical attributes and find women who look something like Halle Berry. While this may appear to be success, if women who look like Halle Berry are attracted to men who look like Eric Benet, then I have a problem, as I am not as attractive as Eric Benet, and may be ignored or turned down by the women who I have found.

While several dating sites do show seekers the preferences of the matches, manual observation and analysis of the match preferences is required in order to use this information. In order to automate the process of utilizing the preferences of the match, a different search philosophy is presented in this paper.

Inverse Profile Matching

'Inverse Profile Matching' (IPM) is the process of generating match results based on the preferences of the search population rather than the seeker. Conventional matching is done by comparing the profiles in the search population against the preferences of the seeker; in mathematical functional notation, this can be expressed as:

$$match\_results = f(see \ker preferences)$$

Inverse Profile Matching generates the search results by comparing the preferences of the search population against the attributes of the seeker, or:

$$match\_results = f(see \ker attributes)$$

If implemented, Profile Attribute weighting can also be used in this inverse manner, so that the attribute weights of the search population members are used to rate the desirability of the seeker.

Whether or not weighting is used, the search results returned by an IPM search will consist of potential matches whose preferences indicate that they will find the seeker desirable.

In sales terminology, conventional matches return a list of cold prospects, while IPM generates lists of qualified, warm prospects.

So despite my Halle Berry infatuation, even if my IPM search does not return any Halle Berry look-alikes, I can now browse the search results and decide who I want to contact with the foreknowledge that anyone I contact from my IPM search results list is guaranteed to find me desirable! This represents a previously unattainable position of confidence that will do wonders for seekers who are shy and / or have low self-confidence in social situations.

Based on current database architectures, it is expected that Inverse Profile Matching will require several orders of magnitude greater processing power than conventional matching, and thus it may be most appropriate for offline search agents rather than online, real-time searching. The increased resource requirements may also be the basis for offering IPM as a premium (i.e. fee-based) feature in sites that implement it.

Activity Weighting

"What you do speaks so loud I cannot hear what you are saying."

This well-known cliché highlights the common dynamic of humans saying one thing and doing another. It is possible that a potential match may have a profile that indicates one preference, but may be taking actions that indicate different preferences.

If certain user actions taken on the site are logged and treated as heavily weighted profile attributes, then the actual behavior of potential matches can be used to improve the quality of search results.

Keeping with my example, my Halle Berry look-alike potential match expresses a preference for Eric Benet look-alikes in her profile, but as there are not many Eric Benet look-alikes contacting her, she has done her own searching and has viewed profiles and sent messages to men that look like me. Without the insight of her actual behavior, her perceived preferences (both to the system and to any potential matches who view her profile) are based more on her ideal (or fantasy) rather than her actual behavior.

If the dating system architecture is designed to support tracking user actions and can model them as profile data, this technique should be a powerful predictor of online dating success, especially when combined with Inverse Profile Matching.

Conclusion

This paper has presented a new way of analyzing existing and new dating profile data in a novel manner so as to significantly improve the possibilities of finding compatible matches using an online dating service.

These techniques do not appear to currently be in use in any dating sites analyzed thus far, nor have I found any theses, white papers or other documents that propose these techniques, and thus I believe that this is a new business process that is appropriate for protection under the patent laws of the United States of America.