Integrating Recommender Information in Social Ecosystems Decisions

Renato A C Capuruço
University of Western Ontario, Dept of Electrical and Computer Engineering
London, Ontario, CANADA, N6A 5B9
r.capu@uwo.ca

Luiz F Capretz
University of Western Ontario, Dept of Electrical and Computer Engineering
London, Ontario, CANADA, N6A 5B9
lcapretz@eng.uwo.ca

ABSTRACT
The exploration of online social ecosystems whose members share mutual recommendations and interactions is a time-dependent and contextual-based process which aims to predict the social status among them. To address the difficulties associated with the process, this article presents the integration of the predictive recommender, social networks, and interaction components into a single methodology. The originality of the proposed framework stems from developing each model based on: (1) a time history and decay algorithm to consider temporal recommendations and interactions; (2) a predictive-aggregating function for different types of social contexts; and, (3) a homophily algorithm to evaluate people's interconnections proximity. Details of the framework are described, a recommender search strategy methodology integrating all of the above is devised, and a case study is used to demonstrate its capabilities. Possible extensions are then outlined.

Keywords
Social Ecosystems, Social Networks, Recommender Systems

1. INTRODUCTION
Over the years, there has been many collaborative recommender systems developed in academia and industry, such as *GroupLens* ([11],[12]) and *VideoRecommender* [5], as well as the book recommendation system for *Amazon.com* and *Metacritic.com* for suggesting movies. These methodologies are largely rooted under principles of aiding individuals to make informed guesses about what other artifacts or items they may also like. More recently, a wealth of social-based information derived from online communities of practice has become increasingly available. In online environments, such as *WebMD.com* for supporting those who seek health information and *canadian-lawyers.ca* for information and resources on legal matters, people interact with each other to build upon the social spaces, casual interactions, and meaningful exchanges that occur.

This differs from the mainstream environments in which recommender systems have born, long-lived and flourished because within social ecosystems settings like the above, there are many other new factors which can influence a person in making personalized choices, especially when doing so upon other people. For example, the *homophilous* phenomenon as described in [9] affects personal judgment, which is influenced by characteristics like age, gender, race/ethnicity and education ([1], [8]), and by psychological characteristics like intelligence, attitudes, and aspirations [13]. Human cognition processes are also influenced by the social contexts people are embedded in and the position of importance they occupy in the network [15]. As such, not only it is important to consider dependencies over time among social contexts where recommendations made in one context predict recommendations in another context, but also to assess and quantify interrelations that create opportunities to achieve societal gain. Ideally, one would like to model as many of these factors as possible in a socially networked recommender system.

However, such conditions complicate and render traditional methodologies as insufficient to deal with the distinct formulation involved. In order to provide a superior exploration and discovery of interconnected social spaces based on individual preferences, this paper proposes a framework that integrates recommender information into social ecosystems decisions more comprehensively. This paper is focused on the architectural aspects of the formulations rather than specific performance features.

2. PROPOSED FRAMEWORK
The main components of a social ecosystem recommender model (SERM) that incorporates a time-dependent, interaction-aware, and social context-sensitive modules are (Figure 1):

- Detailed SERM models (computer-mediated interaction, social network relationships, and recommendation-based perceptions of taste or opinions)
- SERM Constraints (social context, social relation, social recommendation, user defined constraints, such as time horizon, decay intensity, etc.)
- SERM Decision Support Module (user interface, community database, social closeness assessment, recommender search, visualization)

2.1 Model Components
At the core of a successful SERM are proper models for eliciting the interactions and capturing the relationship patterns of the individuals in the community. This allows the several recommendation statements to be evaluated according to the network that is formed, which helps derive recommender-aware social status indicators. As a result, the benefits of knowing those patterns and conditions can be returned to the community itself by

ECSA 2010, August 23-26, 2010, Copenhagen, Denmark.

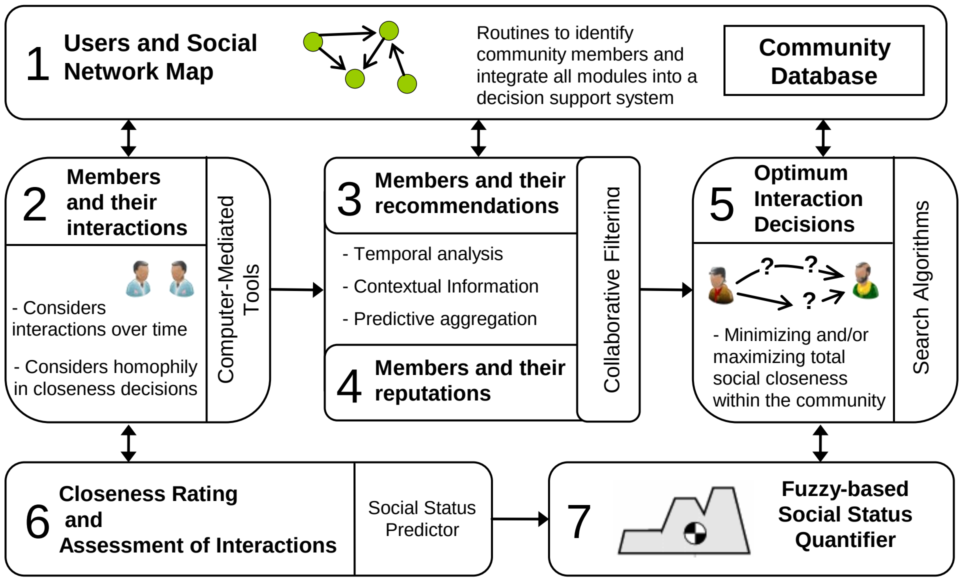


Figure 1. Components of the proposed model.

the means of a search engine that optimally reaches a member from any interrelated member.

2.1.1 Recommender Model

This model is composed by uniquely identifiable members that express opinions of taste and rating values on other uniquely identifiable members. All the information may or not be made public to every other member and every member is able to express them whenever they prefer. Members express recommendations values in members based on their perceived quality as source of advise about past and current topics of discussion. For example, a member should recommend another member if she/he *likes* one's opinions, behavior, contribution, etc. to the overall development of the community.

Every member can express one or more *recommendation statements* that embed his/her opinions about the likability of another member anytime in any social dimension. Recommendations statements have the following form:

- Recommendation(FromMember,ToMember).value
- Recommendation(FromMember,ToMember).date
- Recommendation(FromMember,ToMember).context

Every member's recommendation statement can be formalized in a recommender function whose domain is *M* and whose co-domain is [1, 5] where 1 means total *dislike* and 5 total *like*. A missing value (i.e., function not defined) represents the fact that the member has not expressed a recommendation statement about another member. This is as follows:

$$r_{[\text{member}_m,\ \text{date}_d,\ \text{context}_c]} : \mathrm{M} \rightarrow [1, 5] \cup \perp \qquad (1)$$

For example, $r_{[25,\ 39448,\ 3]}(m_3) = 4.55$ means that member m_{25} issued a recommender statement on *01/Jan/2008* (39,447 days after January 1, 1900) and rated member m_3 as *4.55* in the social context 3 (e.g., workplace), a very high recommendation rating expressing his/her almost complete likeability for the member.

2.1.2 Interaction Model

This model is composed by uniquely identifiable members that express contact with other uniquely identifiable members. All the information may or not be made public to every other member and every member is able to make contact with any other member whenever they prefer. Members express mutual relationships with members based on their actual communication patterns within past and current topics of discussion. For instance, a member should contact (e.g., though e-mailing, discussion board posting, blogging, etc.) another member if she/he *interconnects* with one's opinions, behavior, contribution, etc.

Every member can express one or more *interaction statements* that embed his/her contacts with another member anytime in any social dimension. Recommendations statements have the following form:

- Interaction(FromMember,ToMember).value
- Interaction(FromMember,ToMember).date

Every member's interaction statement can be formalized in an interaction function whose domain is also *M* and whose co-domain is [*x*] where *x* is a calculated value that measures the social distance (i.e., *Closeness*) between two members and it is a function of the Social Status that each member shares in the network. This is explained in more detail in the next subsections. A missing value (i.e., function not defined) represents the fact that the member has not expressed an interaction statement with another member (e.g., "members have not met"). This is as follows:

$$i_{[\text{member}_m,\ \text{date}_d]} : \mathrm{M} \rightarrow [\mathrm{x}] \cup \perp \qquad (2)$$

For example, $i_{[33,\ 39062]}(m_6) = 2.78$ means that member m_{33} issued an interaction statement on *12/Dec/2006* (39,062 days after January 1, 1900) and the calculated social distance to member m_6 is *2.78*. The order of magnitude of the social distance's calculated value is relative to the recommender function's co-domain boundaries because *Closeness* is calculated from the recommendations; therefore, inheriting its scale.

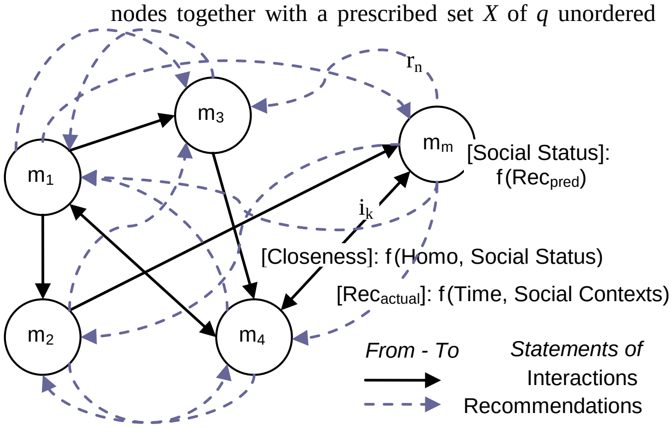
Interactions are considered to be the main avenue upon which recommendations are passed over from one member to another. The matter of fact is that, in this model, it takes one and only one formal interaction between a pair of members to broadcast all (un)known statements of recommendations given over time and in different social contexts by the community. These two aspects are very important features to consider because it not only allows capturing local recommendations (e.g., from members and their immediate neighbors), but also global opinions (e.g., from anyone to everyone else); therefore, bringing all the community together to validate all of the recommendations.

2.1.3 Social Network Model

A social networks perspective is based on the theoretical constructs of sociology and mathematical foundations of graph theory. Classic research in "*sociograms*" and "*sociometric*" [10] established the typical analyses and mathematical models [4] that are used today to understand and analyze social network data.

Based on these generalizations about the features of personal networks combined with the previously introduced formulations for interactions and recommendations, the basic concepts upon which this model is constructed has emerged as follows:

- A graph G consists of a finite nonempty set $V=V(G)$ of p nodes together with a prescribed set X of q unordered



  pairs of distinct nodes of V. Each pair $x=\{u,v\}$ of nodes in X is an edge e of G and x is said to join u and v;

The proposed social network N is a graph where each node p is a member m. Each pair $x=\{m_u,m_v\}$ of members in N is an edge e of N and x is said to join m_u and m_v; a numerical value, $f(e)$, is assigned to each edge e, which is a measure of *Social Closeness*. The complete environment is composed by (Figure 2):

Figure 2. Web of interactions and recommendations representation

- A set M of m uniquely identifiable members:

$$M = \{m_1, m_2, m_3 \dots m_{m-1}, m_m\}$$

- A set R of n uniquely identifiable recommendation statements:

$$R = \{r_1, r_2, r_3 \dots r_{n-1}, r_n\}$$

- A set I of k uniquely identifiable interaction statements:

$$I = \{i_1, i_2, i_3 \dots i_{k-1}, i_k\}$$

- A set C of j characteristics associated with each member m of M:

$$C = \{c_1, c_2, c_3 \dots c_{j-1}, c_j\}$$

- A *Social Status* function $s(m)$ associated with each member m of M:

$$S(m) = f(R_{predicted})_m$$

- A *Social Distance* function $d(i)$ associated with each interaction i of I from members m_a to m_b of M:

$$D(i)_{a,b} = f(S(m)_a, S(m)_b)$$

- A *Closeness* function $c(i)$ associated with each interaction i of I from members m_a to m_b of M:

$$C(i)_{a,b} = f(D(i)_{a,b})$$

2.1.4 Closeness Assessment Model

The assessment model aims at supporting three difficult decisions related to social network exploration, each lending itself well to a different solution mechanism.

First, the process of considering network activities that are dynamic and time-dependent in nature. As such, a time history and decay scheme can aid in the process of harmonizing them to the present time.

Second, unknown localized (between any given pair of members) and a global (among all members) recommendation statements is an issue that may worsen the calculation of a particular's member social status, and as such, improving the accuracy of those calculations is a problem that involves prediction and lends itself well to collaborative filtering application.

Third, the people's social network and implications to the information they receive, the attitude they form, and the interactions they experience sets the stage and contexts for the formation of social spaces in which homophilous relations form and flourish. This is a difficult problem that lends itself well to the consideration of a wide range of socio-demographic and behavioral dimensions to account for the impact of multiplex ties on the dynamics of the network change over time.

Integrating these three components together derives a 5-step calculation. These five calculation steps form the Closeness assessment model with each component / sub-model dealing with one of the sub-problems using a different technique.

2.1.4.1 Time History and Decay

The main goal of the time history and decay sub-model is to account for recommendations expressed in different points in time. It calculates a *decay factor* which accounts for the decreasing effect of importance of a statement of recommendation in different time spans. The proposed model is set up according to the intuition that older statements of recommendation worth less than newer expressed ones.

More precisely, according to the general definitions given by Cohen and Strauss [2] for decay functions, consider a stream of recommendations where $f(t) \geq 0$ is the recommendation value of the stream obtained at time t. For sake of simplicity, it is assumed that the stream only receives values at discrete times, and therefore, t is integral. A *decay function* $g(x) \geq 0$ defined for $x \geq 0$

is a non-increasing function. At time *T*, the weight of the item that arrived at time $t \leq T$ is *g(T-t)* and the *decayed value* is *f(t)g(T-t)*. From that, it is obtained the *decayed sum* of *f(t)* under the decay function *g(x)* that is defined as follows:

$$V_g(T) = \sum f(t)g(T-t), \perp t \leq T \tag{3}$$

A generalized representation of the above formulation is as follows (Figure 3):

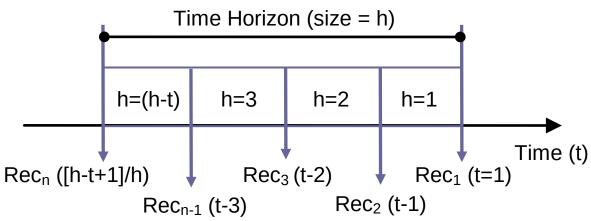


Figure 3. Time history and decay algorithm representation

Assuming *n* recommendations are to be considered up to a maximum time horizon *h*, a recommendation at time *t* from the most recent recommendation will have a decayed recommender factor of (h-t+1)/h. Recommendations that are not reachable within the maximum time horizon have no decayed value.

As an example, let's suppose a member *m* has had 9 recommendations in the course of 5 years of being in a community. Considering a time horizon t=4years, let's say that only 4 of those 9 statements fall within this horizon, one per year: in this case, each of the 4 affected recommender metrics would be multiplied by factors of (4-1+1)/4=1 for the most recent, (4-2+1)/4=0.75 for the 2^{nd} most recent, (4-3+1)/4=0.50 for the 3^{rd} most recent, and (4-4+1)/4=0.25 for the oldest one, respectively.

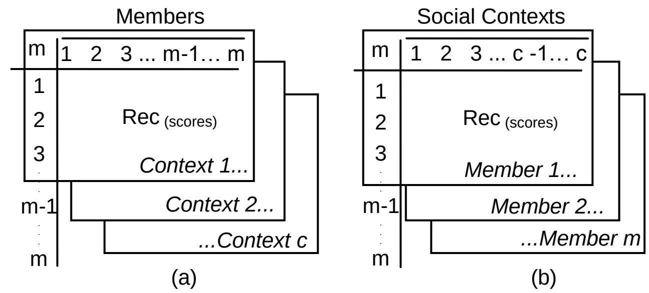
In this way, a linear decay propagation function is adopted: newest member's recommendations have proportionally higher importance than older ones in accordance with the number of recommendations in the *time horizon*.

2.1.4.2 Predictive Aggregation

The predictive aggregation sub-model is made necessary for two main reasons: (1) to reduce the dimensionality of the many recommendations to a singular aggregated value that can be used for further processing, and (2) to enhance the calculation of the *Social Status* function by predicting missing recommendations.

First, the aggregation calculation engine takes as input the recommendation matrix (representing all the community recommender statements) and produces, as output, an equal matrix of pondered recommendations. This is preformed by using either of two approaches: member- and context-centric (Figure 4).

The member-centric approach takes as input the recommender network as a $M \times M$ matrix where the recommender value *r* on each cell *i*, *j* (if present) represents the recommender rating from member m_i to m_j. Because every recommender statement in the matrix refers to a certain context *c* only, *n* input matrices are generated, one for each context. Next, each context is assigned a degree of importance *w* (weight) and the pondered rating *r'* of member *i* to member *j* is the weighted sum of the ratings in each context *c*. More precisely:



$$r'_{i,j} = \sum_{i=1}^{m} \sum_{j=1}^{m} \left(\frac{\sum_{c=1}^{n} \left(r_{i,j} \times w_c\right)_c}{\sum_{c=1}^{n} w_c} \right) \tag{4}$$

As such, a single output matrix *M* x *M* is produced. Recommendations that are unknown for all social contexts remain unidentified and are excellent candidates for prediction.

Figure 4. (a)Member- and (b) Context-centric prediction approaches

The context-centric approach takes as input the recommender network as a $M \times C$ matrix where the recommender value *r* on each cell *i*, *j* (if present) represents the recommender rating received by member m_i in the context c_j. Because every recommender statement in the matrix refers from a certain member *m* only, *n* input matrices are generated, one for each member. Next, the pondered rating *r'* of member *i* in context *j* is the simple average of the ratings received from each member *m*. More precisely:

$$r'_{i,j} = \sum_{i=1}^{m} \sum_{j=1}^{c} \left(\frac{\sum_{m=1}^{n} \left(r_{i,j}\right)_m}{n} \right) \tag{5}$$

As a result, a single output matrix $M \times C$ is produced. Recommendations that were not received from all members remain unknown and are excellent candidates for prediction.

Either of the previous outcome matrices is the traditional input to a Collaborative Filtering (CF) algorithm whose main function is to predict the missing recommendation values. If the member-centric approach matrix is to be used, it lends itself very well to compute missing recommendations against every other member. On the other hand, the context-centric approach matrix is mainly geared towards the computation of missing recommendations taking into consideration dependences among different contexts. This is achieved by using CF's classical steps, as follows:

- *Similarity Metric*: The goal is to calculate the correlation of two overlapping members (represented as vectors of ratings), outputting a $m \times m$ Member Similarity matrix in which *i*th row contains the similarity values of *i*th member against every other member. Pearson's correlation coefficient (6) is the most used technique, as follows:

$$p_{a,u} = \frac{\sum_{i=1}^{m}\left(r_{a,i} - \bar{r}_a\right)\left(r_{u,i} - \bar{r}_u\right)}{\sqrt{\sum_{i=1}^{m}\left(r_{a,i} - \bar{r}_a\right)^2 \sum_{i=1}^{m}\left(r_{u,i} - \bar{r}_u\right)^2}} \quad (6)$$

Both positive and negative similarities values are calculated because similar and dissimilar members *u* to the current member *a* are important measures to grasp the overall community feeling about *a* and, therefore, cannot be ignored.

- *Rating Predictor*: The predicted recommendation rate of member *i* for the current member *a* is the weighted sum of the ratings given to member *i* by the *k* neighbors of *a*. This is the classical CF's last step, as follows:

$$r_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{k} p_{a,u}\left(r_{u,i} - \bar{r}_u\right)}{\sum_{u=1}^{k} p_{a,u}} \quad (7)$$

2.1.4.3 Social Status

Once all missing recommendations have been predicted, it is necessary to compute the *Social Status* for each member. If the member-centric approach was used, hence producing a matrix of size $M \times M$, the Social Status s_a for each member m_a is the simple average of the ratings given to this member by all of the other members (except itself), as follows:

$$s(m_a) = \frac{\sum_{i=1}^{m} r_{a,i}}{m-1} \quad \perp \quad a \neq i \quad (8)$$

Alternatively, if the context-centric approach was used, hence producing a matrix of size $M \times C$, each context *c* in *n* number of social contexts is assigned a degree of importance *w* (weight). Next, the Social Status s_a for each member m_a is the weighted sum of the ratings in each context *c*, as follows:

$$s(m_a) = \frac{\sum_{c=1}^{n}\left(r_{a,c} \times w_c\right)}{\sum_{c=1}^{n} w_c} \quad (9)$$

2.1.4.4 Social Distance

The social distance *d* represents the perceived strength of the relationship between a pair of members. It is a direct function of *Social Status*, and as such, its computation simply averages out the predicted Social Statuses s_i and s_j of a pair of interconnected members m_i and m_j, respectively. This is shown as follows:

$$d_{i,j} = \frac{s(m_i) + s(m_j)}{2} \quad (10)$$

2.1.4.5 Homophily Computation

The main goal of the homophily sub-model is to adjust the social distance "*Closeness*" between a pair of members based on their reciprocal interaction and similarity of personality attributes. Defined by Lazarsfeld and Merton [7], the homophily theory states that most human communication will occur between a source and a receiver who are alike. Homophily implies that distance in terms of social characteristics translates into network distance, the strength of relationships (i.e., interactions) through which a piece of information (i.e. recommendation) must travel to connect two individuals.

More specifically, the homophily computation takes as input the calculated Social Distance *d* from member m_i to m_j and assigns a degree of importance *w* (weight) to each of their matching homophily feature *h* of *n* available features. Next, an overall homophily coefficient *c* is calculated as a product of each w_h to represent the extent by which each original social distance *s* from member *i* to *j* should be shortened. Then, the adjusted distance *d'* from member *i* to member *j* is $d \times c$. More precisely:

$$d'_{i,j} = \left(\prod_{h=1}^{n} w_h\right) \times d_{i,j} \quad \perp \quad h_i = h_j \quad (11)$$

As an example, considering that two members in a community have a calculated social distance of 2.34 and only two matching features from the set *C* of characteristics, each being assigned the importance of 50% and 80%. Then, the social distance would be calculated by the weighted sum of those, that is, 2.34 × (0.5 × 0.8) = 0.936.

In this way, a homophily computation function is implemented where a given original social distance between two members is reduced by a degree that is equal to the combined effects of each matching homophily feature. Unmatched features between two members have no value; therefore, are excluded from the calculations.

3. PROTOTYPE AND VALIDATION

The proposed model was implemented on a commercial spreadsheet program. In this study, Microsoft Excel software is selected for the implementation of the proposed model because of its ease of use and share, and powerful programming features.

Using the Visual Basic language of Microsoft Excel, various procedures were coded to form a complete temporal- and interaction-aware application. These developments involved a substantial effort in coding the several components and providing a user interface.

Since no work in the literature can be used for comparison purposes in terms of the temporal recommender model and social quantifier herein formulated, a case study based on new data is used to demonstrate the capabilities of the developed framework, which is explained in the upcoming section.

3.1 Prototype capabilities: A case study

In order to better understand the current strengths and weaknesses of the proposed methodology, a case study was conducted. The case study used data that was derived upon investigation of an open community of practice called Eyeknowledge.net. This is a virtual community where a variety of people from the eye care industry comes together to share their various knowledge in the topic though online discussions. Each one of the members shares a public profile with socio-demographic, areas of interest and specialties information. In addition, there are many online discussions on several related themes which include a voting system where each one of them may elicit their preference to a member's post. All of these features helped deriving data for the interactions and recommendation statements from one member to another over time, the social contexts (i.e., themes) in which they were made. The data in Table 1 shows the homophily parameters entered into the model for each member. Table 2 represents the derived dimensions in which conversation among members occur, thus was used as social contexts.

As per the proposed model's specifications, a complete social network based on the above was inputted into the system for experimentation. A small random sample of 35 members and 74 statements of interaction among them were used for demonstration purposes. In addition, 5,000 other complete recommendation statements were generated, imported and readily made available to the model. The much larger number of recommendations is to account for the several social contexts and a 5-year recommendation period.

Table 1. List of members' profile data

No.	Attribute	Value
1	Gender	Male, Female.
2	Age	0 to 20 (Children/Youth), 20 to 34 (Younger Adults), 35 to 49 (Adults), 50 to 65 (Older Adults), Over 65 (Seniors).
3	Education	Secondary or less, Technical/Trade training, Post secondary (college, university), Post graduate (Masters and above).
4	Role	Administrative Staff, Industry, MD, OD, Optician, Researcher, Student/Resident, Technical Staff, Patient.
5	Interest	Business, Clinical, Medical.
6	Specialty	Cataract, Contact Lenses, Cornea, Cornea and External Disease, Equipment, Glaucoma, Industry, Neuro-ophthalmology, Ophthalmic Pathology, Ophthalmic plastic surgery, Pediatric Ophthalmology, Refractive Surgery, Retina, Retinal physiology and pathology, Vitreoretinal Diseases, Others.

Table 2. List of social contexts

No.	Attribute	Description
1	Lifestyle	Any matters related to how to improve personal health and well-being
2	Drugs	encompass all topics related to use of prescription drugs, their effects and related (mis)conducts
3	Prognosis	any forecast about the course or outcome of an illness
4	Diagnosis	any opinions derived from the process of identifying or determining the nature and cause of a disease or injury through evaluation of patient history, examination, and review of laboratory data
5	Treatment	expressed opinions on necessary care provided to improve a medical condition, procedures or applications that are intended to relieve illness or injury
6	Business	outlooks on services available to commercial clients who offer assistance with marketing, brand awareness, as well as, providing guidance relating to techniques for treating various ocular disorders

Once the network is defined, the model is ready for verification and experimentation, as described in the following subsection.

3.2 Recommender Search

The primary goal of the recommender search technique is to guide a source member to a destination member through only the most recommended and closest members, not any members. In order to achieve that, the *search algorithm* built upon a Path-based optimization algorithm [3] identifies all members, analyses their interactions, and constructs a social cost structure among them all based on their social closeness.

Implementing a graph-based search algorithm in the framework involves four main steps: (1) eliciting community members, their interactions and recommendations in a given period of time and social contexts; (2) setting the source and destination members; (3) deciding on the evaluation criteria, that is, higher or lower *Closeness*; and, (4) applying the algorithm's relaxation principle to generate short paths from one member to another whose total cost is the least (or most) among them all, to finally display the list of members ones should follow to optimally reach the destination member.

Because the path algorithm is natively design to search for shortest-paths only, the search procedure was adapted to also find paths with highest costs (i.e., less recommended), thus providing greater capability. Being the cost structure a function of the *Social Closeness index* with lower values indicating closer (i.e., more recommended) members, by transforming the scale in which the "short" distance was originally calculated, finding "distant" people is made possible. More precisely:

New Closeness = (High Closeness +1) – Original Closeness (12)

As an example, in a scale from 1 to 5, where 1 represents closer members and 5 designates more distant people, 2 is changed to 4 and 4 is changed to 2; therefore, allowing any more distant members to be also considered by the shortest-path algorithm. After defining the cost structure, the constraints considered in the algorithm are (Figure 5):

a. Choose one or more Social Contexts, assigning corresponding degrees of importance – This will filter out members, interactions and recommendations that are not of interest;

b. Choose one or more homophily parameters, assigning corresponding weights – This will affect the calculation of the final Social Closeness value for each pair of member;

c. Choose time horizon and decay frequency – This concerns to the importance that old recommendations should have as compared to newer ones, ultimately affecting the combined recommender value corresponding for the whole period

d. Specify whether to use lower or higher closeness – This will affect the selection of preferred intermediate members to reach the desired destination member.

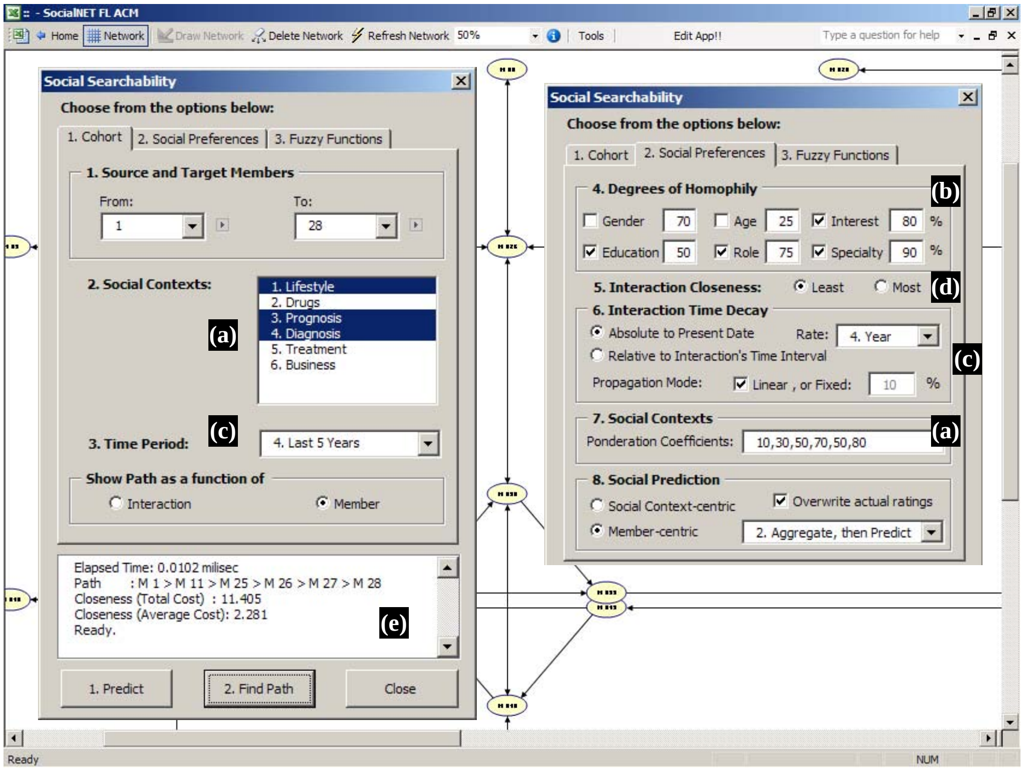


Figure 5. User-defined search strategies and network interface

As an output, a possible solution that gives the list of members (Figure 5.e) within that criterion is presented to the user.

4. DISCUSSION

The developed framework has been demonstrated to work successfully on the example application presented in this article. Various other problems with different combinations of personal networks with different properties and different conditions were experimented with, and the system performed well.

There are a number of possible extensions to the model currently being pursued and caveats being addressed, including:

- *Decay Functions*: The proposed model assumed that recommendations proportionally weaken to the number of periods in the *time horizon*. This is the linear decay model; however, there are many other types of decay functions such as exponential and polynomial where the proportionality of the decay varies with time as well. Experimenting with different types of decay functions to evaluate their effects in the model's output is needed;
- *Time History*: It can be defined in any time unit, such as years, months, weeks, days, etc… which is a very powerful added feature because the more periods *h* in the time horizon, the less sensitive the decay will be in respect to time.
- *Homophily Computation*: due to the lack of standardization, other formulation schemes could be sought. As described, the present model employs a so-called "All-or-Nothing" concept where the social distances between a pair of members is always shortened by a certain amount; however, it could have employed a so-called "80-20" concept where matched features would shorten social distances by the larger number amount while unmatched features would lengthen them by the remaining. This could lead for more precise estimation of *Closeness*;
- *Social Distance Calculation*: presently, the model simply averages out each of a pair of member's Social Status; it could be the case, however, to develop and test other formulations that could potentially lead to different results. For example, a strategy so-called "lowest-wins" could be devised in which the Social distance between two connected members would be the lowest social status of the two; conversely, the "highest-wins" strategy would consider the greater of the two;
- *Search Algorithms*: The choice to use *Dijkstra's* was mainly because it is a well-known, broadly accepted and flexible algorithm to implement despite its limitations (e.g., the algorithm will fail for negative *Closeness*). Because of that, more experimentation with other classes of algorithms is needed, which could not only improve performance (by cutting down on the size of the sub-network that must be explored), but also lead to different results;
- *Collaborative Filtering*: Cold-Start and data sparseness are well-known phenomena in the research literature that could hinder the effectiveness of Recommender Systems' prediction. The implications of these in the model were not considered.

For this demonstration, the size of the network was kept small to be manageable. However, more tests with real size online communities with thousands of people are welcome to extend the methodology. Other enhancements that could be made include migrating the framework to a more advanced web-based interface with improved visualization capabilities.

4.1 Remaining Developments

Fuzzy-based social status quantifier. Working together with the assessment model, a promising mechanism that uses linguistic variables to handle the vagueness and imprecision in determining the relative importance that a member occupies in the network is being sought to integrate reputation information into the current model. The development of this component uses the concept of fuzzy-set theory originated by Zadeh [6] and the concepts of fuzzy control developed by Takagi–Sugeno [14]. In recommender-aware social network decisions, fuzzy linguistic variables such as "recommender rating" (*Rec*), "reputation rating" (*Rep*), and "closeness distance" (*C*) are fuzzy variables that represents the social status and social distance between any pair of members, respectively. The linguistic variables can be represented by a family of linguistic terms (fuzzy sets *VL, L, M, H, and VH*) that covers the space of "recommendation-rating" solutions ranging from "very low" for VL to "very high" for VH, for instance. Then, these preconditions have to be stored in the form of rules (called *fuzzy rules*) along with the decision maker's preference in their associated social status and closeness rating. An example rule is *"IF Recommendation (Rec) is high (H) AND Reputation (Rep) is low (L) THEN Closeness (C) is far (F)"*.

5. CONCLUSION

In this article, a recommender model was developed incorporating time-dependent and social-aware recommendations, an interaction-based social network quantifier to identify the proximity of their members, and a modified search algorithm to optimally reach members. All framework components proved to work efficiently in support of network exploration, discovery and reach decisions.

The recommender model stores rated opinions of one member to another over time and in different social contexts. These help in predicting unknown opinions more truthfully, as more factors are considered. The social network and their quantifiers, on the other hand, effectively translates the opinions of taste into the networks' closeness relationships that account for the interaction flow, improved guidance for social status and social distance computations, and user preference of interacting with similar or dissimilar members. Lastly, the search algorithm proved to work successfully for practical size problems to find members close or apart from each other. To facilitate use of the model by practitioners, the model was implemented as a prototype spreadsheet system that is easy to use and share. The prototype allows the user to insert, delete and update any number of members, their interactions and recommendation statements, draws the network, and automates the search optimization. A case study was used to demonstrate the capabilities of the system.

6. REFERENCES

[1] Bott, H. (1928). "Observation of play activities in a nursery school". Genetic Psychology Monographs, 4:44-88.

[2] Cohen, E., Strauss, M. (2006). "Maintaining Time-Decaying Stream Aggregates". Journal of Algorithms. Volume 59, 1:19-36, Elsevier Press.0

[3] Dijkstra, E. W., (1959. "A Note on Two Problems in Connexion with Graphs". In Numerische Mathematik, 1, S. 269–271.

[4] Harary F.,(1969). "Graph Theory", Addison-Wesley Publishing Company, page: 9-10.

[5] Hill, W., Stead, L., Rosenstein, M., and Furnas, G. (1995). "Recommending and evaluating choices in a virtual community of use". In Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems, 194 – 201

[6] L. A. Zadeh (1965). "Fuzzy sets". Information and Control 8 (3) 338–353.

[7] Lazarsfeld P., MertonR.K., (1954). "Friendship as Social Process: A Substantive and Methodological Analysis", In Freedom and Control in Modern Society, 18-66, ed. M.Berger and C.Page, New York, Octagon Books.

[8] Loomis, C. P., (1946). "Political and Occupational Cleavages in a Hanoverian Village". Sociometry. 9, 316-333.

[9] McPherson, M., L. Smith-Lovin, and J. Cook (2001). "Birds of a feather: Homophily in social networks". Annual Review of Sociology 27, 415-44

[10] Moreno, J.L., (1951). "Sociometry: Experimental Method and the Science of Society – An Approach to a New Political Orientation" Beacon House, Beacon, New York.

[11] Resnick, P., Iakovou, N., Sushak, M., Bergstrom, P., and J. Riedl. (1994). "GroupLens: An open architecture for collaborative filtering of netnews". In Proceedings of the 1994 Computer Supported Cooperative Work Conference

[12] Resnick, P., Varian, H., (1997). "Recommender Systems", Introduction to Special Section of Communications of the ACM, vol. 40(3):56-58.

[13] Richardson, H. M., (1940). "Community of Values as a Factor in Friendship of College and Adult Women". Journal of Social Psychology, 11:303-312.

[14] Sugeno, M.; Kang, G.T., (1988), "Structure identification of fuzzy model", Fuzzy Sets and Systems, 28, pp. 15-34

[15] Wasserman, S., Galaskiewicz, J. (1994) "Advances in Social Network Analysis. Research in the Social and Behavioral Sciences". SAGE Focus Editions. 4:92-104