Published in th. Journal of Network and Computer Application, Elsevier 2007,
doi:10.1016/j.jnca.2007.04.004 (Available Online May 2007)

Computing in Social Networks with Relationship Algebra

Javed I. Khan & Sajid S.Shaikh
Internetworking and Media Communications Research Laboratories
Department of Computer Science
Kent State University
233 MSB, Kent, OH 44242, USA
Javed|sshaikh@kent.edu

Abstract

Communities are the latest phenomena on the Internet. At the heart of each community lies a social network. In this paper, we show a generalized framework to understand and reason in social networks. Previously, researchers have attempted to use inference-specific type of relationships. In this paper, we propose a framework to represent and reason with general case of social relationship network in a formal way. We call it relationship algebra. In this paper, we first present this algebra then show how this algebra can be used for various interesting computing on a social network weaved in the virtual communities. We show applications such as determining reviewers in a semi-professional network maintained by conference management systems, finding conflict of interest in a publication system, or to infer various trust relationships in a community of close associates etc. We also show how future community networks can be used to determine who should be immunized in the case of a contagious disease outbreak and how these networks could be used in crime prevention etc.

Key words: Social network, relationship, trust, conflict-of-interest, reviewer selection, immunization, data mining, web computing.

1. Introduction

Emerging Era of Social Networks: Communities and social networking are one of the latest and fastest growing phenomena on the Internet. The websites providing social networking services are fast becoming an important cog in the borderless world of Internet. One can view them as digital town squares where different kind of people having varied interests can interact with each other.

As the web increasingly becomes an influential part of people's lives, the distinction between the actual and the virtual social network is rapidly fading. One can know a great deal about a person even without physically meeting or talking to him. An individual is progressively depending upon these virtual social networks for companionship, advice, entertainment, education etc. Communities such as Orkut®[8], Yahoo360®, MySpace®, LinkedIn® are providing a platform for such social interactions. Orkut is a website designed specifically for friends and family. The whole thrust of Orkut is to make the conversation with friends and family more upbeat and fun. It further allows members to create communities, so that like-minded people can come together and have lively and engaging discussions. Orkut's use as a social tool is complex, because various people frequently try to add strangers to their own pool of friends, more often than not just to increase the number indicating their number of friends in their profile. LinkedIn is an example of another social networking service geared towards professionals. Even though the primary objective of many of these websites is to connect people over the Internet, the audience they serve is often dissimilar to each other and they offer variants of privacy settings and communication tools. Table-1 lists some of the leading community networks of today. Fig-1 provides a snapshot portal from two sites. These portals represent just one node in a vast network comprised of millions of nodes. The graphics identifies the link types provided by these services. The database of various conference and editorial management systems can also be viewed as example of business network which pioneered explicitly collecting relationship among authors, program committee volunteers, reviewers, and editors while they engaged in conference organizations. Many of these services themselves provide a host of social network powered communication tools to the community.

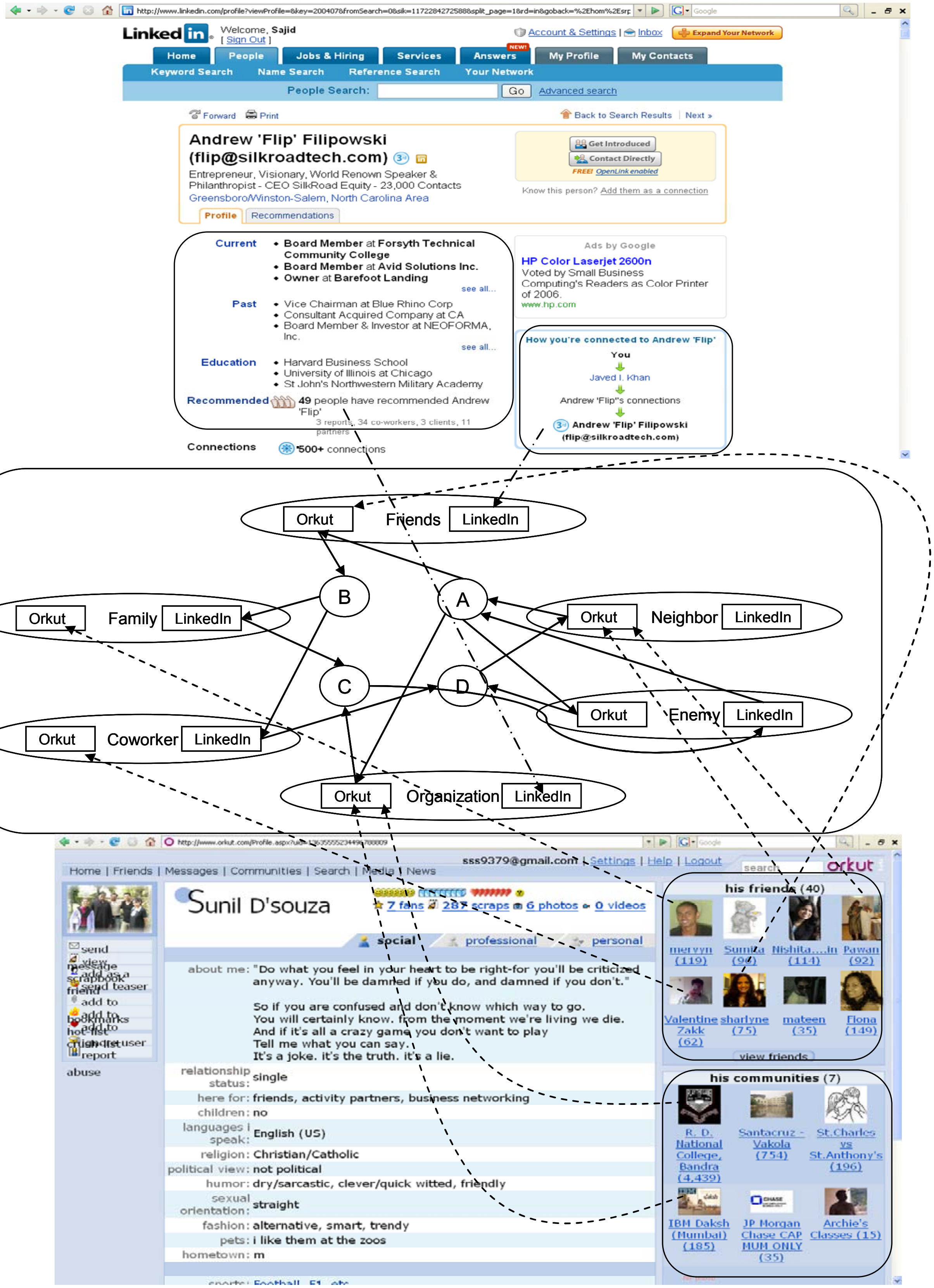

Figure 1: The top part shows a LinkedIn® web page, the middle part shows a social networks schema graph and the bottom part of the figure shows an Orkut ® web page.

Services, such as Google's Gmail®, have structured their expansion of social network. Also various peer-to-peer networks are weaved in the fabric of social network. At the very heart of these systems is an extensive relationship network. Very powerful applications are conceivable from the global relationship information available in them.

Social Network Studies of the Past and Present: Formal study of social networks can be traced back as early 1930's. Time and often it has been a fascinating topic of interests among social scientists. The early emphasis however was to understand dynamics of relationship. The social network and the interactions, which take place among them, have been studied in [1, 2]. In another work, different flavors of friendships between individuals have been studied in [3, 4]. [5, 6] analyzed the business relationships between companies. These studies used technique for limited type of relationships and limited propagation. The problem with pre-web social network study was the availability of data. It was very difficult to obtain the actual relationship network. In [9] the author study the flow of information from one person to other in his social circle. In [10] the researchers have analyzed the voting patterns of the US senators and have found that senators from the same state tend to vote similarly.

The advent of web opened up access to larger social network data. Google® search can also be considered as some form of social network system, which tried to analyze importance or relevance of papers, based on links placed in web pages by users in a networked web. The reader/writer relation considered these as a form of recommendation. Later more direct recommender systems such as Epinion® emerged which would take the recommendations in more explicit terms but would employ more Eigen computing to decide or infer the final strength of relationship. One interesting study in relationship computation is the study of trust propagation. Guha and Kumar [7] presented a framework for the propagation of trust in a directed network of people. The scholarly publication management systems and citation networks maintained by various networked services perhaps are the first-broad based public attempt to maintain explicit network. What makes the virtual communities distinct is that now these are poised to more explicitly encode and use social network relationships.

Use of Social Networks: It is very likely that such wealth of information about an individual could be used in very powerful way. Using the social network profile on Orkut and one's scrapbook one can make a calculated guess about the relationships between the individual and people in his friend circle. The scrapbook can also help one to determine the strength of these relationships. If one wants to find out the set of close friends of his boss, one can easily achieve this through the information available on Orkut. Similarly, if one wants to know one's bosses interests the community section on his profile can give him some sort of starting point. Thus even though this information seems to be very naïve, a concerted social network computing can help to derive relationships which are not obvious on surface but do exist. Similarly, LinkedIn can be used to find out the firms and organizations a person has been affiliated in the past. This information can be used to find a set of people who are likely to have an influence on the individual.

A whole new range of very powerful social network applications –which can be called *society applications* are conceivable based on the social network assets growing under these communities. In this paper, we expose how such a network can be used in a generalized way for various computations on social networks. In section-2 we first show the *relationship algebra* to represent the edges connecting the nodes and the set operations are used to extract and infer complex relationship between the nodes of the network for deriving relationship synthesis and propagation. Then in section 3 we provide several examples of social network based computing. In addition, we specifically show algorithms for conflict of interest determination. We show how it can be user for reviewer selection, panel selection etc. We also show how networks can be used in crime-watch, immunization planning, programmable trust determination, etc. The relationship algebra can be used for aggressive applications which can determine political decision making, means to influence personal decision, as well as expose complex business connections.

Table 1: Survey of Some Current Social Networks

Network	Members	User Base
Orkut	22,000,000	Designed specifically for friends and family.
LinkedIn	6,000,000	Designed for professionals and adults.
MySpace	54,000,000	Used mostly for fun and blogging.
Sporzoo	2,000,000	Real Estate Investors and Professionals.
SelectedMinds	1,000,000	Corporate social networking.

2. Relationship Algebra

2.1. Representation

The world is comprised of unique entities. Each unique entity (E) is represented by an entity ID i.e. Entities in this world are however, organized as members of various sets. There are sets such as author (A), paper

(P), journals (J), reviewer (R), etc. Each set has members, which are entities. An entity can be member of multiple sets. For example, individual 'Andrew' can be a member of an author set as well as of reviewer set. Members have also membership index in each set. The membership index of an entity does not have to be the same between sets. In a way, all objects in this world are members of the super set E. In this world–, various pairs of sets can have relationship. For example, papers have authors, i.e., set A and set P have relationship *author-to-paper*. Thus, a member in set A may or may not be an author-to-paper relation with each member of the paper set P.

Let A is a set (vector) of members of set author, and P is a set (vector) of members of type papers, then the cross product M^r=AXP is the matrix of *author-to-paper* relationship, we call it *relation matrix*. Each element mij represents the strength of relationship. In binary valued strength if m_{ij}=1 it represents individual a_i has an author-to-paper relationship to paper p_j, or m_{ij}=0 indicates the absence of this relationship between the two. However, m_{ij} can have other values as well.

We will use the notation $M_{i,-}$ to present the i-the row of matrix M. It is a relationship statement about i-th member of A and says who in P are related with i-th member. We will use the notation $M_{I-,j}$ to present the j-the column of matrix M. It is a relationship statement about j-th member of set P and says who in A are related to this j-th member. Two sets can have more than one kind of relationship, each represented by a separate relation matrix M. We the superscript r to denote the connecting sets and the specific relationship.

2.2. Relationship and Set Algebra

Now, we define a set of operations on relation matrices. If A is a relationship matrix then we define following semantic operators:

(i)Equivalence (two relations are semantically same: biological brother's biological father is biological father), (ii)synthesis (matrix product: two ordered relationship to produce the notion of a more complex relationship, example: father's brother is an uncle),(iii)reflection (matrix transpose: reflexive inverse of a relation: example husband: wife),(iv)absence (element wise 1-mij: friend: 1-friend) on the matrix,(v) semantic inverse (semantically no functional form is known to computer the relation matrices: example friend vs. enemy, trust vs. distrust).

We also define the fowling set operations: intersection (element wise AND), dediagonalization (all diagonal set to zero, the intension is to remove self-to-self relations), set exclusion (element wise subtract), set union (element wise OR), quantization (such as all elements is mapped to zero or one for binary quantization).

We introduce the following notations to denote the above operations. Equivalence is denoted by A=B, synthesis by A x B, reflection by A^T, absence by $\overline{A}$, semantic inverse by $\tilde{A}$, Following notations are used for element to element operations on relationship matrices: intersection of two relations by $A \otimes B$, union of two relations by $A \oplus B$, exclusion of related set B from related set A by AΘB, dediagonalization $\hat{A}$, and quantization $\lceil A \rceil$. The above operations set now enables us to define, track, infer, and analyze various complex social relationships, and their interplays.

Table 2: Set Operations

Operation	Sym.	Explanation
Relation Source Set Extraction	Ψ	$\Psi_i^X(M)$ where $X\ is\ M_{ij} \geq \mu$ or $X\ is\ M_{ij} < \mu$ Determines the column set for each row M_i. This is the set consisting of all M_j such that $M_{ij} \geq \mu$
Relation Target Set Extraction	ρ	$\rho_j^X(M)$ where $X\ is\ M_{ij} \geq \mu$ or $X\ is\ M_{ij} < \mu$ Determines the row set for each column M_j This is the set consisting of all M_i such that $M_{ij} \geq \mu$
Winner Target	ξ	$\xi_j^X(M)$ where $X\ is\ M_{ij} > \forall M_{kj}$ Determines the row M_i for M_j such that M_{ij} is highest value in the column M_j
Winner Source	ϕ	$\phi_i^X(M)$ where $X\ is\ M_{ij} > \forall M_{ik}$ Determines the column M_j for row M_i such that M_{ij} is highest value among in the row M_i
Zero Column	θ	$\theta^X(M)$ where $X\ is\ M_{ij} = 0$ and $0 < i < n$ Determines each column M_j such that all the elements in that column at zero.

Table 3: Relationship Operations

Operation	Symbol	Explanation
Equivalence	A=B	$a_{ij} = b_{ij}$
Reflection	R =A^T	$r_{ij} = a_{ji}$

Synthesis	S = AxB	$s_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$
Intersection	E=A⊗B	$e_{ij} \; where: 0 < i < n, 0 < j < m = \left[a_{ij} \cap b_{ij}\right]$
Union	U=A⊕B	$u_{ij} \; where: 0 < i < n, 0 < j < m = \left[a_{ij} \cup b_{ij}\right]$
Exclusion	X= AΘB	$x_{ij} \; where: 0 < i < n, 0 < j < m = \left[a_{ij} - b_{ij}\right]$
Dediagonalization	$\left[\hat{A}\right]$	$A_{ij} \; where: 0 < i < n, 0 < j < m = \left[0 \; if \; i == j\right]$
Quantization	$\left[A\right]^{\mu}$	$a_{ij} \; where: 0 < i < n, 0 < j < m = \begin{bmatrix} 1 \; if \; a_{ij} \geq \mu \\ 0 \; if \; a_{ij} < \mu \end{bmatrix}$

3. Examples of social networks

3.1. Schema Graph of a Publication Network

The major entities involved in the publication network are the Authors, the Organizations, the Paper, the Journal, the Reviewers, the Editors and the Topic Area. The topic area is the focal point of the network. It is evident from the schema graph presented in figure 2 that there exist direct relationships between various nodes of the graph. These are visible relationships, our aim is to find the ones, which are not obvious but have a profound affect on the publication decisions. For example, the relationship between the author of a paper and the editorial board of the journal in which the paper has been submitted for publication. We expose these relationships using the example graph of a publication network shown in figure 6 of Appendix A. We would be using the instance graph for illustrating how we can use the relationship algebra presented in section 2 for useful purposes such as "*Finding a set of reviewers for a particular paper"*. The only relationships available to us are the ones indicated by the arrows going from the source to the sink. For example, the arrow drawn from "*Jeroen Dietz"* to *"*P_1*"* represents an Author → Paper relationship. The author → organization relationship can have many flavors such as student, professor, boss, employee, and researcher. This is because the organizations can be varied, such as universities, research labs, private companies, government funded establishments etc. Using the primary relationships enumerated in the table 4 and applying the relationship algebra upon them enables us in deriving and detecting interesting phenomenon such as conflict of interest.

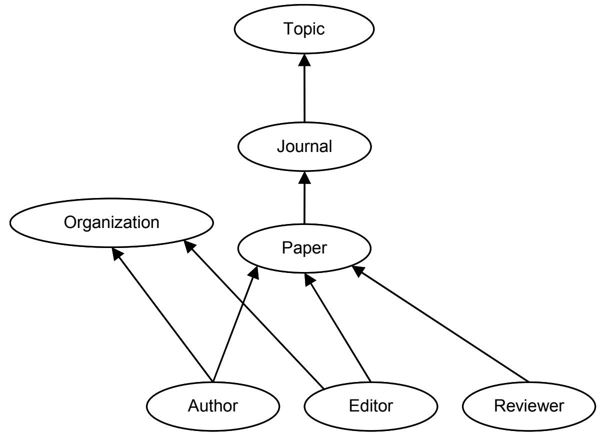

Figure 2 :Schema Graph of Publication Network

3.2. Application: Reviewer Selection

The network can be used to select a set of authors who can be on the reviewing committee of a paper such that they meet certain restrictions. The reviewer selection can be expressed by a set of constraints. Below is an example set:

Reviewer Selection Constraints: *(i)The reviewer should not be a coauthor of the paper he is going to review, (ii) he should not be a coworker of the author for example the author and the reviewer should not be faculties in the same university. (iii) The reviewer should not have submitted a paper in the same journal or conference and (iv) finally he should be well acquainted with the subject area being discussed in the paper.*

Table 4 Primary Relationship for The Publication Network

Primary Relationship	Notation
Journal → Topic Area	$M^{J-T}_{J_i}$
Editor → Paper	$M^{E-P}_{E_i}$
Paper → Journal	$M^{P-J}_{P_i}$
Author → Paper	$M^{A-P}_{A_i}$
Reviewer → Paper	$M^{R-P}_{R_i}$
Author → Organization	$M^{A-O}_{A_i}$

We have illustrated the complete reviewer selection process by working out an example shown in figure 8 of Appendix B. The aim of the example is to find a reviewer set for papers P_5 and P_6 from among the four authors available. The first step is to determine the authors and coauthors for P_5 and P_6 . This is achieved by multiplying

Published in th. Journal of Network and Computer Application, Elsevier 2007,
doi:10.1016/j.jnca.2007.04.004 (Available Online May 2007)

the matrix M_A^{A-P},which represents the relationships between the authors and the papers with its transpose matrix. The resultant matrix $M_{coAuthor}$ represents the co-author relationship between the respective authors. In the next step we determine, which authors have submitted papers in the same journal. In order to determine this we first need to establish a relationship matrix depicting the relationship between the authors and the journals M_A^{A-J} , which is done by multiplying the matrices, M_A^{A-P} and the one representing papers-journals relationships M_P^{P-J} . The resultant co-journal matrix $M_{coJournal}$ is a product of M_A^{A-J} and its transpose matrix. Now we have to determine, which all authors are coworkers. The coworker matrix $M_{coWorker}$ is computed by multiplying the matrix representing the authors-organization relationships M_A^{A-Org} and its transpose. Finally we determine the non conflict of interest matrix $M_{nonConflict}$ by subtracting each of the *coAuthor, coJournal* and *coWorker* matrices from the matrix depicting the relationship between all the authors belonging to the same topic area M_{all}. The reviewer set matrix $M_{reviwer}$ is calculated by multiplying the $M_{nonConflict}$ and the $\left(M_A^{A-P}\right)^T$ matrices. Applying the row extraction set operation ρ on the reviewer set matrix gives us the reviewer set for papers P_5 and P_6.

$$M_{reviewer} = M_A^{A-P}\left[M_{all}\,\theta\, M_{\text{coAuthor}}\,\theta\, M_{\text{coJournal}}\,\theta\, M_{\text{coWorker}}\right] \quad ...(1)$$

$$\text{Re}\ viewerSet(P_i) = \rho_i^{M_{ij}=1}\left(M_{reviewer}\right) \quad ...(2)$$

In the above discussion we have mentioned the term *conflict of interest*, which can be defined as follows. A **conflict of interest** consists of three entities, the source *"i"*, the sink *"j"* and the relationship between them *"R"*. It occurs if we have two distinct relationship trails R_1 and R_2 from *i* to *j* and their intersection set is nonempty.

$$S_i^j(R_1) \cap S_i^j(R_2) \neq \phi \quad \ldots(3)$$

We can determine the *conflictset* for each author in the above example by applying the column extraction set operation Ψ on the reviewer set matrix $M_{nonConflict}$.

3.3. Application: Panel Selection

Another use of the network can be in the selection of intellectuals for the formation of a panel for a particular research area. The panel should satisfy the following constraints. For our example the constraint set is given below.

Panel Selection Constraints*:(i)The members of the panel should consist of people from different fields of the area and (ii)they should belong to varied organizations such as universities, research labs, industry etc.*

The algorithm for panel selection is as follows. The first step is to extract the expert in each field from the M_A^{A-J} matrix using the max column set operation "ξ" to form the *expertset*. Then for each panelist in the extracted set we determine the kind of organization represented by him or her through the M_A^{A-Org} matrix. The zero column set operation is applied to the M_A^{A-Org} matrix to ascertain that all organizations have been represented on the panel. If the operation results in a nonempty set, then a person from the missing organization is picked from the M_A^{A-Org} matrix using row extraction to give the *missingexpert* set. The panel is the union of *expertset* and *missingexpert* sets.

$$\text{expertset}\left(T\right) = \xi_j\left(M_A^{A-J}\right) \quad ..(4)$$

$$\text{missingorg}\left(\text{T}\right) = \theta\left(M_{A_i}^{A-O}\right) \quad ..(5)$$

$$\text{missingexp ert}\left(\text{T}\right) = \rho_{\text{missingorg}}^{\text{M ij}>0}\left(M_{A_i}^{A-O}\right) \quad ..(6)$$

$$\text{panelset}\,(T) = \text{expertset}\,(T) \cup \text{missingexp ert}(T) \quad ..(7)$$

3.4. Schema Graph of a Social Network

An individual's social network primarily consists of family, friends, neighbors, coworkers and the organizations with which he is affiliated. The circles denoted by A, B, C and D are the individuals in a community. The ellipses denote the six major types of relationships we have considered for our example. The smaller rectangles within the ellipses denote the refined relationships for each class. These refined relationships are enumerated in the table below.

Table 5: Refined Social Relationships

Relationship Class	Relationship Objects
Family	Father, Son, Daughter, Spouse
Friend	Good Friend, Acquaintance
Enemy	Competitors, Different Beliefs
Coworker	Boss, Colleague, Subordinate, Partner
Neighbor	Next Door, Same Community
Organization	Educational, Religious, Entertainment, Philosophical.

The instance graph shows the social network of an individual *"George"*. *George's* social network consists of his family, his fellow workers, his neighbors, the organizations he is associated with and his enemies. Each of these nodes further has their own social networks, which are a part of *George's* network, but *George* has a derived relationship with the nodes of these secondary networks. The strength of *George's* derived relationships depends upon the strength of his primary relationships.

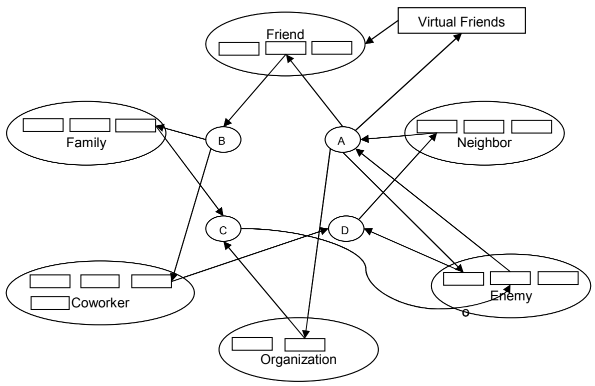

Figure 3: Schema Graph of Social Network

Table 6: Primary Relationship for The Social Network

Primary Relationship	Notation
Individual → Company (Owner)	$M_{Ind}^{Ind-Company(Owner)}$
Individual → Friend	$M_{Ind}^{Ind-Ind(Friend)}$
Individual → Father	$M_{Ind}^{Ind-Ind(Father)}$
Individual → Org (Member)	$M_{Ind}^{Ind-Org}$
Individual → Neighbor	$M_{Ind}^{Ind-Ind(Neighbor)}$
Individual → Enemy	$M_{Ind}^{Ind-Ind(Enemy)}$
Individual → Boss	$M_{Ind}^{Ind-Ind(Boss)}$
Individual → Coworker	$M_{Ind}^{Ind-Ind(Coworker)}$
Individual → Spouse	$M_{Ind}^{Ind-Ind(Spouse)}$
Individual → Org (Client)	$M_{Ind}^{Ind-Org(Client)}$

3.5. Application: Immunization

Suppose George is a virus carrier and we want to find out the people who might have been infected by him and needs vaccination. In order to achieve this we need to determine the *vaccination set* from George's social network. The people to be included in the set should satisfy certain conditions. For our particular example, the conditions are as follows:

Immunization Constraints: *(i) The people most vulnerable are the ones which come in physical contact with George. These are usually friends, family, neighbors and coworkers. They have to be immunized.*(ii) *The second group of people who are likely to get infected are the ones which belong to George's derived network i.e. his greater than 1 hop neighbors. The likelihood of them been infected depends upon their relationship strength with George's 1 hop neighbors. For our example, the threshold value is 0.6*

Using these two relationship strengths, we compute the matrices $M_{George}^{Spouse}, M_{George}^{Father}, M_{George}^{Neighbor}, M_{George}^{Coworker}, M_{George}^{Friend}$, which each represent the derived relationship strength between George and his greater than one hop neighbors. Then we apply column extraction for the first row of each of these matrices to get an individual subset.

$$set\left(A\right)^{Spouse} = \psi_{A}^{M_{ij}>0.6}\left(M_{A}^{Spouse}\right) \quad ..(8)$$

The final vaccination set (V) is a union of all the individual subsets.

$$V\left(A\right) = set\left(A\right)^{Spouse} \cup set\left(A\right)^{Father} \cup set\left(A\right)^{Friend} \cup set\left(A\right)^{Coworker} \cup set\left(A\right)^{Neighbor} \cup set\left(A\right)^{Son} \quad ..(9)$$

3.6. Application: Crime-Watch

The network can be used for crime prevention. Using the network information the police can get together a surveillance team that would help to keep a watch on the places likely to be visited by a fugitive. The fugitive will almost certainly receive help from his family members and friends. The constraints for becoming a member of the surveillance team are as follows.

Surveillance Team Constraints: *(i) The team member should a neighbor of either fugitive's family members or his friends. (ii) He should not be a friend of the fugitive's family or fugitive's friends.*

The surveillance set can be found by determining the relationship matrices between the fugitive say "*George*" and the neighbors of his family and friends. The matrices $M_{George}^{Neighbor(Friends)}$ $M_{George}^{Neighbor(Family)}$ represent the relationship between George and the neighborhood of his family and friends. The surveillance matrix $M_{George}^{Survillance}$ is the union of these two matrices, we exclude the neighbors who are the fugitive's friends, which are given by the matrices $M_{George}^{Friend(Family)}$ and $M_{George}^{Friend(Friend)}$ .The surveillance set is obtained from the matrix using column extraction

Published in th. Journal of Network and Computer Application, Elsevier 2007, doi:10.1016/j.jnca.2007.04.004 (Available Online May 2007)

$$M_{George}^{Survillance} = \left\{ M_{George}^{Neighbor(Family)} \oplus M_{George}^{Neighbor(Friends)} \right\} \theta M_{George}^{Friend(Family)} \theta M_{George}^{Friend(Friend)} \quad (10)$$

$$\text{survillanc eset}(\text{George}) = \psi_{George}^{M_{ij}=1}\left(M_{George}^{Surveillance}\right) \quad ..(11)$$

The social network of George can be used to determine which people and organizations have influence on him. This information is very important if one wants to manipulate his decision on certain matter such that it benefits ones interest. The influential set would contain entities, which have a strong relationship with George such as friends, family, church and the ones, which could affect his finances such as business partners, boss, and banks. The first set of individuals can be easily determined since they are direct relationships personal(George). The second set business (A) is determined from the matrix $M_{George}^{Partners}$, which represents his business relationships.

$$\text{influences et}(A) = \text{personal}(A) \cup \text{business}(A) \quad ..(12)$$

3.7. Application: Trust Propagation

As human dependence on the internet as a source of reference before making an important decision increases, there is growing need to differentiate between trustful and distrustful sources. Most of the time we cannot determine it by ourselves, but have to infer it from the experiences of people in our social network. Researchers [7] have studied how to infer various forms of trust in a complex relationship network. They have suggested four atomic trust propagation techniques viz. direct propagation, co-citation, transpose trust and trust coupling.

The relationship algebra can be used to define the above mentioned propagation techniques as well as various other forms of trusts and also determine various combinations and synthesis in a programmable way. There are various ways is which trust propagation can be achieved and each individual has a choice as to which path he would take to determine trust. This is because everybody has a different notion of trust and may not share the same principles as someone else does in determining trust.

An example of various trust relations are shown in the social network instance graph of figure 4. In figure 9 in Appendix B we have worked out an example for illustrating a few of the trust propagation techniques. If *George* trusts *Laura* represented by matrix M_{spouse} and if *Laura* trusts *Peter* is represented by matrix M_{Friend} then the product of these two matrices M_{result} shows that *George* trusts *Peter*. This is an example of *transitive propagation*. If *George* trusts *Laura, Jenna* is represented by matrix M_A and *Bernard* trusts *Laura* is represented by matrix M_B then the product of M_A ,its transpose and M_B shows that *Bernard* trusts *Jenna*. This is an example of *inferential propagation*. If *George* trusts *Laura* is represented by M_{spouse} then a product of M_{spouse} and its transpose shows that *Laura* also trusts *George*. This is an example of *reflexive propagation*. If *George* trusts *Jenna* , is represented by M_{father} matrix and *Jenna* trusts *Kent State University* and *Lily* trusts *Kent State University* are represented by M_{org} then the product of the three matrices M_{father}, M_{org} and transpose of M_{org} shows that *George* trusts *Lily*. This is an example of trust union propagation. A complete workout of this example is illustrated in figure 9.

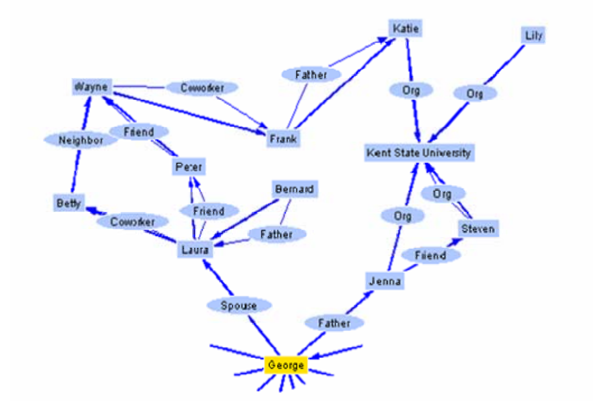

Figure 4: Instance graph used to demonstrate trust propagation

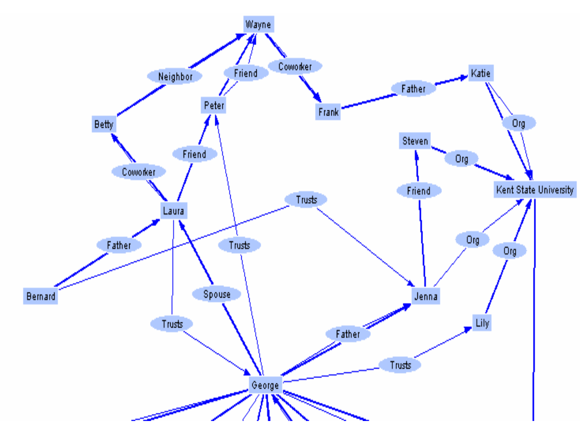

Figure 5: Result graph demonstrating trust propagation

4. Conclusions

Over the last couple of years, a number of social networking websites have become the cornerstones for social interactions. The information being shared by individuals on these websites is becoming richer by the day. In this paper, we have presented a framework, which can be applied to these social network and the information thus obtained can be used to construct a whole new range of applications. We have used two social networks viz. a publication network and a social network, to illustrate the application of the relationship algebra on networks. In the process, we have also proposed a few applications such as reviewer selection, panel selection, immunization and crime watch.

The algebra helps in deriving complex and apparently hidden relationships in almost algebraic manner, which may not be obvious to any individual owner of information in the chain.

The computation in social systems, unlike relational databases inherently requires real valued computation such as relationships strength propagations. The instance graphs tend to be unbounded. Therefore, we have presented the algebra using simple matrix centric operation- which allows localization and suitable propagation of strength values. Multi-valued logic (such as Fuzzy Logic) might be applicable in social computing. Extensions of Codd's relation algebra [11] itself is an ongoing topic of investigation [12, 13, 14].

The paper provides just a glimpse at how seeming simple information shared by daily users can be used in a very powerful way by the social networking services. It raises valid concern about the boundary of information ownership and privacy issues. In the light of such developments, it is important to be aware of the overwhelming implications in privacy.

5. References

[1] Scott, J., "*Social Network Analysis: A Handbook"*, Sage Publications, London, 2nd ed. (2000).

[2] Wasserman, S. and Faust, K., "*Social Network Analysis"*, Cambridge University Press, Cambridge (1994).

[3] Moreno, J. L., "*Who Shall Survive?",* Beacon House, Beacon, NY (1934).

[4] Rapoport, A. and Horvath, W. J., *"A study of a large sociogram"*, Behavioral Science 6, 279–291 (1961).

[5] Mariolis, P., *"Interlocking directorates and control of corporations: The theory of bank control"*, Social Science Quarterly 56, 425–439 (1975).

[6] Mizruchi, M. S., "*The American Corporate Network*, 1904–1974, Sage, Beverley Hills (1982).

[7] R. Guha, Ravi Kumar et.al. "*Propagation of Trust and Distrust".* ACM WWW2004 New York, New York, USA 2004.

[8] Meneses, J., "*The Orkut.com case: a reflection on the exploration of new ways to online sociability in the tradition of the study of virtual communities".* (2004)

[9] Fang, W., Bernardo, A., et.al., " *Information Flow in Social Groups"*. (2003)

[10] Jakulin, A., Buntine, W., "*Analyzing the US Senate in 2003: Similarities, Networks, Clusters and Blocs".* (2003)

[11] Codd, E.F., "*A Relational Model of Data for Large Shared Data Banks*". Communications of the ACM 13 (6): 377–387.(1970)

[12] Buckles, B.P., Petry, F.E., "*A fuzzy representation of data for relational databases.*" Fuzzy Sets Systems. 7:213–226, (1982).

[13] Abiteboul, S., et.al., "*The Lowell database research self-assessment.*" Comm. ACM 48(5):111–118.(2005)

[14] Radim Belohlavek, Vilem Vychodil, *"Codd's Relational Model of Data and Fuzzy Logic: Comparisons, Observations, and Some New Results," cimca*, p. 70, International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06), 2006.

[15] Radim Belohlavek. "*Fuzzy Relational Systems: Foundationsand Principles"*. Kluwer, Academic/Plenum Publishers, NewYork, 2002.

[16] Raju., K. V. S. V. N., Majumdar., A. K." *Fuzzy functional dependencies and lossless join decomposition of fuzzy relational database systems*". ACM Trans. Database Systems Vol.13, No. 2:129–166, 1988.

Published in th. Journal of Network and Computer Application, Elsevier 2007,
doi:10.1016/j.jnca.2007.04.004 (Available Online May 2007)

6. Appendix A – Instance Graphs

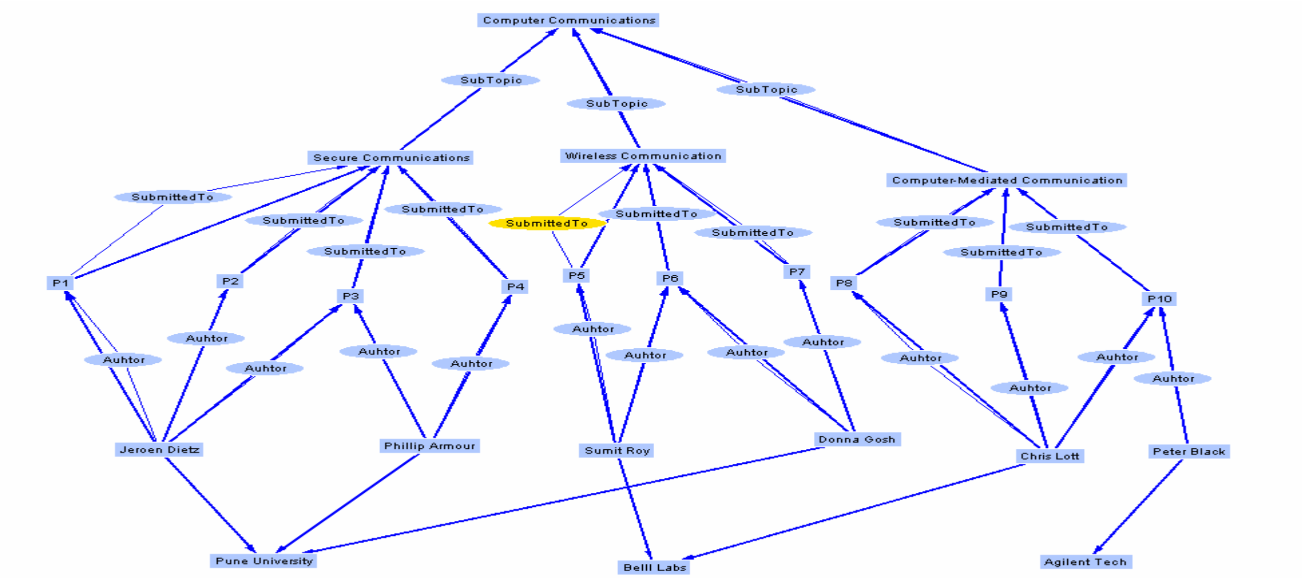

Figure 6: Instance graph for the publication network

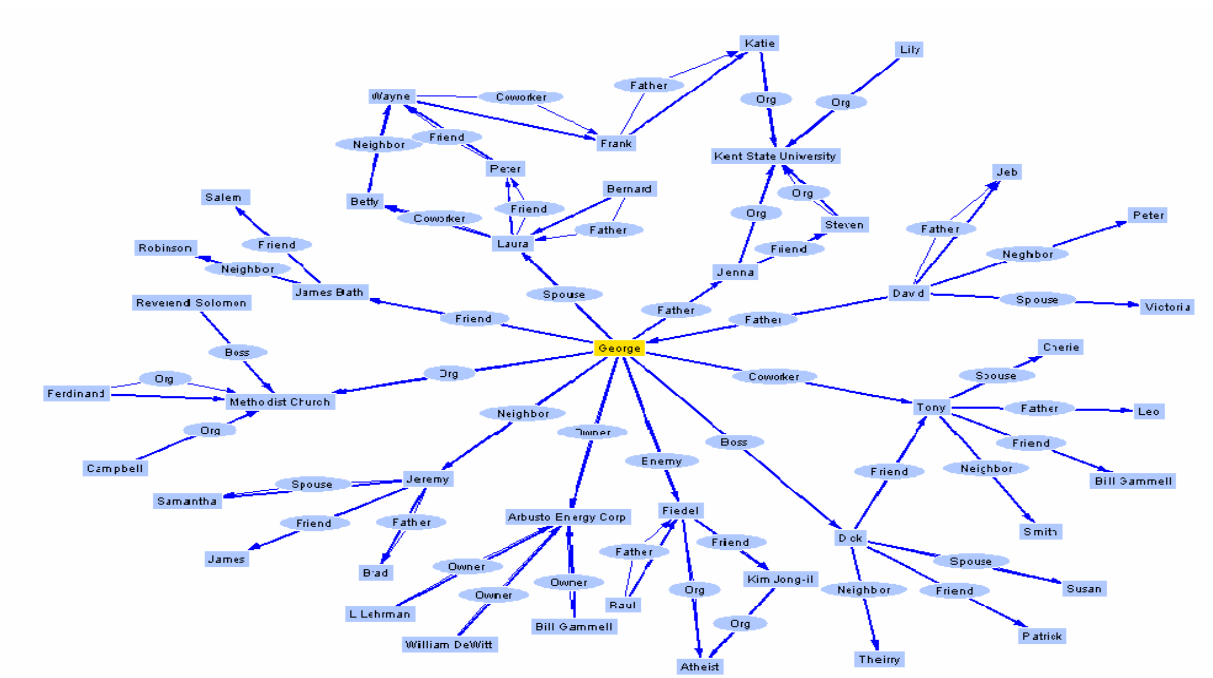

Figure 7: Instance graph for Social Network

Published in th. Journal of Network and Computer Application, Elsevier 2007,
doi:10.1016/j.jnca.2007.04.004 (Available Online May 2007)

7. Appendix B - Examples

Finding A Reviewer Set for papers P_5 and P_6

Step 1 :Derive The Co-Author Matrix $M_{coAuhtor} = M_A^{A-P} \times \left(M_A^{A-P}\right)^T$

$$
\underbrace{\begin{array}{c|cc} & P5 & P6 \\ \hline JD & 0 & 0 \\ PA & 0 & 0 \\ SR & 1 & 1 \\ DG & 0 & 1 \end{array}}_{M_A^{A-P}}
\times
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline P5 & 0 & 0 & 1 & 0 \\ P6 & 0 & 0 & 1 & 1 \end{array}}_{\left(M_A^{A-P}\right)^T}
=
\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 0 & 0 \\ PA & 0 & 0 & 0 & 0 \\ SR & 0 & 0 & 2 & 1 \\ DG & 0 & 0 & 1 & 2 \end{array}
\xrightarrow{[A]^1}
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 0 & 0 \\ PA & 0 & 0 & 0 & 0 \\ SR & 0 & 0 & 1 & 1 \\ DG & 0 & 0 & 1 & 1 \end{array}}_{M_{coAuhtor}}
$$

Step 2 :Derive The Co-Journal Matrix $M_{coJournal} = M_A^{A-J} \times \left(M_A^{A-J}\right)^T$ Where $M_A^{A-J} = M_A^{A-P} \times M_P^{P-J}$

$$
\underbrace{\begin{array}{c|cc} & P5 & P6 \\ \hline JD & 0 & 0 \\ PA & 0 & 0 \\ SR & 1 & 1 \\ DG & 0 & 1 \end{array}}_{M_A^{A-P}}
\times
\underbrace{\begin{array}{c|cc} & SC & WC \\ \hline P5 & 0 & 1 \\ P6 & 0 & 1 \end{array}}_{M_P^{P-J}}
=
\begin{array}{c|cc} & SC & WC \\ \hline JD & 0 & 0 \\ PA & 0 & 0 \\ SR & 0 & 2 \\ DG & 0 & 1 \end{array}
\xrightarrow{[A]^1}
\underbrace{\begin{array}{c|cc} & SC & WC \\ \hline JD & 0 & 0 \\ PA & 0 & 0 \\ SR & 0 & 1 \\ DG & 0 & 1 \end{array}}_{M_A^{A-J}}
\times
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline SC & 1 & 1 & 0 & 0 \\ WC & 0 & 0 & 1 & 1 \end{array}}_{\left(M_A^{A-J}\right)^T}
$$

$$
= \underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 0 & 0 \\ PA & 0 & 0 & 0 & 0 \\ SR & 0 & 0 & 1 & 1 \\ DG & 0 & 0 & 1 & 1 \end{array}}_{M_{coJournal}}
$$

Step 3 :Derive The Co-worker Matrix $M_{coWorker} = M_A^{A-Org} \times \left(M_A^{A-Org}\right)^T$

$$
\underbrace{\begin{array}{c|cc} & \text{Pune University} & \text{Bell Labs} \\ \hline JD & 1 & 0 \\ PA & 1 & 0 \\ SR & 0 & 1 \\ DG & 1 & 0 \end{array}}_{M_A^{A-Org}}
\times
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline \text{Pune University} & 1 & 1 & 0 & 1 \\ \text{Bell Labs} & 0 & 0 & 1 & 0 \end{array}}_{\left(M_A^{A-Org}\right)^T}
=
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 1 & 1 & 0 & 1 \\ PA & 1 & 1 & 0 & 1 \\ SR & 0 & 0 & 1 & 0 \\ DG & 1 & 1 & 0 & 1 \end{array}}_{M_{coWorker}}
$$

Step 4 :Determine the non-conflict matrix $M_{nonconflict} = \left[M_{all} \,\theta\, M_{coAuthor} \,\theta\, M_{coJournal} \,\theta\, M_{coWorker}\right]$

$$
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 1 & 1 & 1 & 1 \\ PA & 1 & 1 & 1 & 1 \\ SR & 1 & 1 & 1 & 1 \\ DG & 1 & 1 & 1 & 1 \end{array}}_{M_{all}}
-
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 0 & 0 \\ PA & 0 & 0 & 0 & 0 \\ SR & 0 & 0 & 1 & 1 \\ DG & 0 & 0 & 1 & 1 \end{array}}_{M_{coAuthor}}
-
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 0 & 0 \\ PA & 0 & 0 & 0 & 0 \\ SR & 0 & 0 & 1 & 1 \\ DG & 0 & 0 & 1 & 1 \end{array}}_{M_{coJournal}}
-
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 1 & 1 & 0 & 1 \\ PA & 1 & 1 & 0 & 1 \\ SR & 0 & 0 & 1 & 0 \\ DG & 1 & 1 & 0 & 1 \end{array}}_{M_{coWorker}}
$$

$$
= \begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 1 & 0 \\ PA & 0 & 0 & 1 & 0 \\ SR & 1 & 1 & -2 & -1 \\ DG & 0 & 0 & -1 & -2 \end{array}
\xrightarrow{[A]^1}
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 1 & 0 \\ PA & 0 & 0 & 1 & 0 \\ SR & 1 & 1 & 0 & 0 \\ DG & 0 & 0 & 0 & 0 \end{array}}_{M_{nonconflict}}
$$

Step 5 :Determine the reviewer matrix for P_5 and P_6 $M_{reviewer} = \left(M_A^{A-P}\right)^T \times M_{nonconflict}$

$$
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline P5 & 0 & 0 & 1 & 0 \\ P6 & 0 & 0 & 1 & 1 \end{array}}_{\left(M_A^{A-P}\right)^T}
\times
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline JD & 0 & 0 & 1 & 0 \\ PA & 0 & 0 & 1 & 0 \\ SR & 1 & 1 & 0 & 0 \\ DG & 0 & 0 & 0 & 0 \end{array}}_{M_{nonconflict}}
=
\underbrace{\begin{array}{c|cccc} & JD & PA & SR & DG \\ \hline P5 & 1 & 1 & 0 & 0 \\ P6 & 1 & 1 & 0 & 0 \end{array}}_{M_{reviewer}}
$$

Figure 8: Reviewer selection example

Published in th. Journal of Network and Computer Application, Elsevier 2007, doi:10.1016/j.jnca.2007.04.004 (Available Online May 2007)

Trust Propagation

1.Transitive Propagation $M_{result} = M_{Spouse} \times M_{Friend}$

George trusts Laura. Laura trusts Peter. So George trusts Peter

M_{spouse}

	George	Laura	Peter
George	1	1	0
Laura	0	1	1
Peter	0	0	1

X

M_{friend}

	George	Laura	Peter
George	1	1	0
Laura	0	1	1
Peter	0	0	1

=

M_{result}

	George	Laura	Peter
George	1	2	1
Laura	0	1	2
Peter	0	0	1

2. Inferential Propagation $M_{result} = M_A \times M_A^T \times M_B$

George trusts *Laura, Jenna*. *Bernard* trusts *Laura . Hence, Bernard* trusts *Jenna*

M_A

	George	Laura	Jenna	Bernard
George	1	1	1	0
Bernard	0	1	0	1

X

M_A^T

	George	Laura
George	1	0
Laura	1	1
Jenna	1	0
Bernard	0	1

X

M_B

	George	Laura	Jenna	Bernard
George	1	1	1	0
Bernard	0	1	0	1

=

M_{result}

	George	Laura	Jenna	Bernard
George	3	4	3	1
Bernard	1	3	1	2

3. Reflexive Propagation $M_{result} = M_{Spouse} \times M_{Spouse}^T$

George trusts *Laura.* Hence *Laura trusts George*

M_{spouse}

	George	Laura
George	1	1
Laura	0	1

X

M_{spouse}^T

	George	Laura
George	1	0
Laura	1	1

=

M_{result}

	George	Laura
George	2	1
Laura	1	1

4. Trust Union Propagation $M_{result} = M_{Father} \times M_{Org} \times M_{Org}^T$

George trusts *Jenna* , *Jenna* trusts *Kent State University .Lily* trusts *Kent State University* Hence, *George* trusts *Lily*

M_{father}

	George	Jenna	Lily
George	1	1	0
Jenna	0	1	0
Lily	0	0	1

X

M_{org}

	George	Jenna	Lily	Kent State University
George	1	1	0	0
Jenna	0	1	0	1
Lily	0	0	1	1

X

M_{org}^T

	George	Jenna	Lily
George	1	0	0
Jenna	1	1	0
Lily	0	0	1
Kent State University	0	1	1

=

M_{result}

	George	Jenna	Lily
George	3	3	1
Jenna	1	2	1
Lily	0	1	2

Figure 9: Trust Propagation example