Community as a New Communication Layer in the Internet

Hideaki Takeda
National Institute of Informatics
and
The Graduate University for Advanced Studies

2-1-2 Hitotsubashi, Chiyoda-ku
101-8430 Tokyo, Japan
E-mail takeda@nii.ac.jp

This work is done with Masaharu Hamasaki, Ohmukai Ikki, and Toru Takahashi.

The Information Flood on the Internet

- The information flood on the Internet
  - The amount of information on the Internet has been increasing
  - How to ensure the quality of information within the enormous amount of information
- The key to escape from the information flood is human relationship
Human relationship for the information flood

- People can act intelligent agents for each other to collect, filter and associate necessary information.
- They can help not only themselves but also other people
  - *Four eyes see more than two.* (三人寄れば文殊の知恵)
  - *Scratch me and I’ll scratch you.* (魚心あれば水心)
- But, it can work only via reliable human network

Human relationship for the information flood

- Only reliable human network can help us to keep quality of information
  - Which do you believe recommendation of TV programs by your friends or TV guide books?
    - The degree of reliability
    - The degree of closeness
    - The degree of shared interests
    - …
- Community or human network is the key to enable information sharing and exchange with quality
Roles of communities for information sharing/exchanging

- Communities or human network can serve as a layer of communication via computer network
- The distinction
  - Personal human network:
    - the graph where a person as a node and a relation between persons as a link
  - Community:
    - the structure upon personal human network

Roles of communities for information sharing/exchanging

- Roles of communities
  - provide channels for information exchange on specific topics
    - Becoming a member of a community is obtaining a channel to send and receive information related on the community
    - E.g., Mailing list
  - work as filtering information
    - Members of communities collect and distribute information specific to some topics either explicitly or implicitly. Once a consensus can be formed in a community, information specific to the topics can be easily collected or selected by members of the community.
  - offer a field for collaboration to produce new information
    - Communication in communities can yield new information
    - E.g., discussion group on BBS
Tasks to realize “better” online communities

- Two directions
  - Make online communities natural like real-world communities
    - Online communities are still by far un-mature in comparison with real-world communities.
  - Exploit characteristics of online communities
    - Reduce real-world constraints
      - Time, Space, etc
    - Add new communication ways
      - Agents, asynchronous communication, etc
  - Both directions are needed
    *Balance is important*

Tasks

- Forming communities
  - How to know relationship among people?
    - Relating people to each other
      *Re-configuration of personal human network*
    - How to form communities from relationship among people?
      - Finding common needs, interests, topics, etc.
      *Finding relationship via WWW bookmarks*
    - Collaborative Scheduling Support System for Conferences
- Facilitating activities in communities
  - How to make communities better?
    - Exploit merits of online communities and compensate their demerits
      *Expressive Media for Online Communities*
    - e-kyoshitsu: Application to Distance Learning for Children
  - How to utilize information in communities?
    - Filtering, extraction, summarization, etc.
      *(...)*
Re-configuration of personal networks by the neighborhood matchmaker method

M. Hamasaki, H. Takeda

Purpose

- Personal network is usually “ad hoc”
  - We may miss better friends nearby
- We need better network
- One Solution:
  - Collect data for all people, then generate the “best” network
    - Disadvantage:
      - Scalability
      - Privacy
- Our approach:
  - Neighborhood Matchmaker Method (NMM)
Neighborhood Matchmaker Method (NMM)

- A iterative approach to optimize the network
- Every node works as a matchmaker for neighborhood nodes to improve the network
- The basic idea
  - In our real life, introducing new friends by the current friends is a practical way to optimize personal networks
    - We can know persons who you have not known before
    - Your friend can filter people for you
- Advantages
  - No need for central servers
  - Applicable to any size of community
  - Less computational cost

Algorithm

- 1. A node calculates connection values between its neighbor nodes
  - We call that node “matchmaker”
- 2. If the matchmaker finds a pair of nodes which has a good enough connection value, it selects this pair for recommendation. The matchmaker introduces both nodes of recommended pair to each other
- 3. The node that receives recommendation decides whether it accepts or not. If it accepts, it adds a path to the recommended node
Sample Networks

- We compare two networks, (2) and (3)

1. initial (generated randomly)
2. converged (NMM)
3. best (central server model)

Results: Cover-Rate w.r.t. Nodes

- The path size is fixed as three times as the node size
- All cases were converged
- The average of cover-rate and the turn of convergence vary with the node size
Results: Average of Convergence Turn

- The number of convergence turn is linearly increased with the node size.
- Computational cost
  - NMM: $O(N)$
  - Central Server Model: $O(N^2)$

Conclusion

- Proposal of optimization of “ad hoc” network
- Good news for the Internet communities
  - No need for central servers
  - Applicable to any size of community
  - Anytime Algorithm
Discovery of Shared Topics Networks among People
A Simple Approach to Find Community Knowledge from WWW Bookmarks

H. Takeda, M. Hamasaki, T. Matsuzuka, Y. Taniguchi

Purpose

- Generation of human network guiding individual information activities
  - An example
    - I want to watch sports programs on TV. What your recommendation?
  - Who and What
- Shared Topics Network among Users (STN)
Our approach

- Combination of manual and automatic methods
  - Identification of topic
    - Use of bookmark files as users’ knowledge
      To overcome knowledge acquisition problem
  - Discovery of inter-topic relations
    - Text analysis to calculate inter-topic relations

Bookmarks as Knowledge

- A bookmark folder
  = A topic interested by the user
- URLs in a bookmark folder
  = Examples of the topic

Shared Topics Network
Procedure to discovery shared topics

Calculation of Similarity among WWW pages

↓

Estimation of Similarity among folders for different users

Calculation of Similarity among WWW pages

- Extract words in texts
- Pick up some top words in frequency
- Find shared words: count how many these words are overlapped in two pages
- If there are enough shared words, these pages are related
Discovery of common topics

- Count pairs of related pages for every pair of folders
- If they are enough related pairs, these two folders are related.

kMedia Interface
**Bookmark with recommendation**

- Two pages are related
  - Each page is a recommended page to the other

**Discovered Shared Topics Network**

![Diagram of discovered shared topics network]
Discovery of topic relations

- Common relations
  - (search, IR), (academia, research-related)
  - similar but words themselves are different
- Un-common relations
  - (Unix, academia)
  - Speciality of the community

Discovery of relationship among people

- What are common topics with others?
- Who is good at this topic?
Experimental Evaluation (1)
Subjects of Experiment

- 12 subjects
- 3 persons from 4 communities (lab. = community)
- Two tasks for subjects
  - Submit their bookmark files
  - Evaluate recommendations generated by STN
- Two types of groups to generate STN
  - In-community: Belongs to the same laboratory
  - Cross-community: Comes from different laboratories

Experimental Evaluation (2)
Items for User Evaluations

- The evaluation ranges from 1 to 5 (5 is the best)

- Page
  - Are the recommended pages similar?

- Folder
  - Are the recommended folders similar?
  - Are the recommended folders useful?

- Person
  - Do you want to contact the person?
  - Do you want to meet the person?
Analysis of Effects of Community (1)
the Relation Between Page and Folder Relevance

- High correlation between folder relevance and page relevance for in-community case

![Graph showing correlation between folder relevance and page relevance](image)

Category Resemblance (1)
Categorization Is Human Relation?

- Human relation can be measured by resemblance of folder structure

\[ C_{ij} = \frac{N_{fij} \times R_{fij}}{N_{p_{ij}}} \]

- \( C_{ij} \): Category resemblance
- \( N_{f_{ij}} \): No. of recommended folders
- \( R_{f_{ij}} \): Folder relevance
- \( N_{p_{ij}} \): No. of recommended pages
Effects of Category Resemblance (2)
Correlation Coefficient of the Parameters to Evaluations of Person

- The category resemblance is the highest of all parameters in this experiment

<table>
<thead>
<tr>
<th>Parameters</th>
<th>To contact</th>
<th>To meet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category Resemblance</td>
<td>0.49</td>
<td>0.55</td>
</tr>
<tr>
<td>Num. of recommended pages</td>
<td>0.42</td>
<td>0.30</td>
</tr>
<tr>
<td>Ave. of page relevance</td>
<td>-0.13</td>
<td>-0.19</td>
</tr>
<tr>
<td>Num. of recommended folders</td>
<td>0.45</td>
<td>0.30</td>
</tr>
<tr>
<td>Ave. of folder relevance</td>
<td>0.38</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Avg. of Evaluation of page          | 0.29       | 0.40    |
Avg. of Evaluation 1 of folder      | 0.28       | 0.32    |
Avg. of Evaluation 2 of folder      | 0.09       | 0.20    |

Effects of Category Resemblance (4)
for Page Recommendation

- Better page recommendation results for new group made from category resemblance (CR)

![Page Recommendation Score Diagram]
Summary

- Proposal of shared topic network to enhance user’s communication
- Proposal of algorithm of discovery of shared topic networks with WWW bookmark files
- Validity of our approach by an experiment
- Proposal of category resemblance as measurement for community effects

Collaborative Scheduling Support System for Conferences
(on-going project)

H. Takeda, M. Hamasaki
In cooperation with Yutaka Matsuo and Takuichi Nishimura
Purpose

- System Aim: Support people to find their friends in a specific group
- Research Theme: Investigate different human networks in the same group
- Three human networks
  - Human network in the activity: *I worked with him*
  - Human network by communication: *I know him*
  - Human network by behavior: *I meet him*
- Scheduling on conferences
  - Plan and Action

System Functions

- Easy-to-use scheduling system for the conference
  - Just add presentations what you want to watch
- Can refer schedules of other people
  - Manually collaborative scheduling
  - Can only see schedules of *who know you*
- Can recommend schedules (*not yet*)
  - Automatically collaborative scheduling
- On-site support of schedules (*not yet*)
  - Small communication device with sensors

Cobit

Takuichi Nishimura, Hideo Itah, Yoshinobu Yamamoto and Hideyuki Nakashima. “A compact battery-less information terminal (CoBIT) for location-based support systems.” In Proceeding of SPIE, number 4863B-12, 2002.
The current status of the system

TelMeA
Show Me What You Mean - Expressive Media for Online Communities

Toru Takahashi, Yasuhiro Katagiri, H. Takeda
Introduction of TelMeA2002

- What is TelMeA2002?
  - TelMeA2002 is an asynchronous community system like bulletin board system (not internet chat system).
  - TelMeA2002 employs character agents as personal conversational media among users.
  - We call such personal agents as personified media.
  - In TelMeA2002 community, users can make messages in combination of full body expressions and pointing to web contents with personified media.
Conversation Process in TelMeA2002

1. Edit a message in terms of a script language
2. Submit the message to the community server
3. The message is accumulated in the conversation log
4. Request for seeing the message from others
5. The required message is downloaded
6. The message is asynchronously enacted

Our Goal

- Is to find pragmatic rules of social and nonverbal interactions
  - Supporting social and nonverbal interactions
  - Archiving the logs of long-term community activities
  - Analyzing usages and effects of nonverbal expressivity

- Make a model of multimodal social interaction
- Calculate social evaluations for involved information
- Summarize or make reutilize the involved information
Challenges 1 - Identification

- Unique embodiment is necessary for quick identification.
  - Because users are represented by their personified media.
- For analysis, however, each personified medium needs to have same set of expressions.
- Making same animations for various personified media puts a heavy load on the development.

- We first focus on an analysis of usage of various type of animation before expanding the grade of identification.
  (56 kinds of animations for all 8 types of personified media)

Challenges 2 - Communication Features

- Personified media should cover all 4 features of human communication
  - Facts – enable through the spoken content
  - Relationship – expressed through the relative spatial distance and position
  - Appeal – expressed through the selection of various performative verbs
  - Self-revelation – communicated through the emotional expression
Challenges 3 - Expressive Repertoire

- Personified media need to cover the entire scale of expression for the believability.
  - 35 performative verbs (explains, agrees, complains, etc.)
  - 48 affective expressions (likes, sadly, worries, etc.)
  - 13 interpersonal attitudes (yes, I know, forgotten, etc.)
  - Direct attention
    - Pointing, interpersonal distances
- Some essential conversational expressions such as glance and nods are less of importance because of nature of asynchronous conversation.

Trial Use: e-教室(e-classroom) Project

- e-教室(e-classroom) Project:
  - Run by NPO
  - Distance learning for children (mainly junior-high school, 12-15yrs)
  - Several classrooms (math, economics, CG, etc)
- TelMeA for e-教室
  - Experimental use of TelMeA
  - Classroom for
    - Leaning “agent” as new technologies by using
    - Communicating to each other (“BBS” for participants)
  - (demo)
The current status of “TelMeA for e-教室”

- Period: c.a. 4 month (2003.1.16-)
- Login users: 64
- Posted users: 24
- Post No.: 297, Post thread No.: 22
Summary

- Information technologies, in particular AI can offer new opportunities for communities
  - Reducing constraints of the real world
    - Time, space, etc
  - New communication ways
    - Knowing new related people, communication via agents etc
- They will change meaning or roles of communities
  - e.g,
    - Very weak communities
    - Quick life cycle of communities
    - Belonging so many communities

Summary

- Challenges
  - Support of life cycle of communities
    - Create, maintain, diverse, merge, disappear
  - Trust
    - Trust is very difficult
    - Trust may be more complicated than the real world