3

Augmented Analytical Exploitation of a Scientific Forum

Xijin Tang[1], Yijun Liu[2], and Wen Zhang[1]

[1] Institute of Systems Science, Academy of Mathematics and Systems Science Chinese Academy of Sciences, Beijing 100080 P.R.China
[2] Institute of Policy and Management, Chinese Academy of Sciences, Beijing 100080 P.R.China
{ xjtang, yijunliu, zhangwen } @amss.ac.cn

Summary. In reality, group work, especially various meetings, exists as a feasible way for people to communicate and collaborate to deal with problems. Various academic meetings and conferences serve as an important part of social process toward scientific knowledge growth. It is significant to understand more about the outcome of those meetings for maintaining the scale of those dialogues and facilitating policy making. In this paper, some augmented analytical methods are applied to a famous scientific forum on frontiers of science and technology in China, Xiangshan Science Conference, to expose some ignored information which is eagerly required by conference organizers, policy makers and researchers. Those methods, such as visualization of expert opinion structure, augmented information support by Web text-mining, clustering of contributed ideas and various analysis about individual's participation, etc. are integrated into a group argumentation environment (GAE), which aims to support divergent group thinking process for emergence of a *ba* for knowledge creation and provide a variety of perspectives towards the concerned topics by those addressed conferencing mining techniques.

3.1 Introduction

Whatever efficiency or effectiveness of those meetings held in daily life, that kind of group work exists as a feasible way for people to share ideas, interests, understandings and achievements about some focused topics and to search and find solutions toward a variety of problems. For scientific researchers, weekly seminars, academic conferences, scientific forums, etc. are usual ways for collaboration, discussion and exchange of experience and practice on the related issues or topics. By P. Thagard's view, scientific knowledge growth consists of the psychological processes of discovery and acceptance, the physical processes involving instruments and experiments, and the social processes of collaboration, communication, and consensus that brought about transformations in knowledge [1]. Thus it is necessary for governmental departments, research organizations, societies, etc. to maintain those group

X. Tang et al.: *Augmented Analytical Exploitation of a Scientific Forum*, Studies in Computational Intelligence (SCI) **123**, 65–79 (2008)
www.springerlink.com

activities to facilitate sociological approaches to the study of science. Lots of efforts had been invested. However, how to evaluate the outcome of those activities regarding those visible or invisible efforts has not been comprehensively studied by organizers in comparison to those direct outputs, such as budgets, publications, participation, etc. after those group activities were completed.

In China, XiangShan Science Conference (XSSC) is the most famous platform for scientific discussions and debates. Similar to Gordon Research Conferences in USA, XSSC is fully supported by government since 1993 and serves as a scientific forum which consists of a series of small-scale academic workshops where a group of scientists working at the frontier of a particular area meet to discuss in depth all aspects of the most advanced topics in the relevant fields and then new plans for research may be incubated. With its excellent academic contents, broad scientific visions and featured operating mechanism, XSSC has gained renowned reputation in open-mindedness and innovation, made important contributions to national science development and exerted a profound impact on the decision making process of the various government departments concerned. Even with a variety of statistic figures about outputs of XSSC, and aggregation of various records about those workshops including lectures, discussions and debates, comments and summaries posted on the conference web site, few studies have been undertaken toward those records to detect more hidden information by quantitative methods.

In this paper, a suite of analytical methods is applied to some explorations from such a knowledge repository where stored active scientists' understandings, wisdoms in scientific research at a context of economic and social reforms in China in recent 20 years. Versatile analyses are undertaken, such as visualization of expert opinion structure, various clustering of contributed opinions, augmented information support by Web text-mining, various measures about participants' contribution and roles in those series workshops, etc. All those analytical tools have been integrated into our developed group argumentation environment (GAE), which aims to support divergent group process for the emergence of a *ba* for idea generation and knowledge creation and provide a variety of perspectives towards the concerned topics by those conferencing mining techniques.

3.2 Exploitation of XSSC as Group Argumentation

Till now, almost 300 workshops across multiple disciplines and with over 10,000 participants of different ages had been held under the name of XSSC. Due to its features in facilitating interdisciplinary discussions, each workshop could be regarded as a group thinking process toward one or some scientific problems. Some hot topics, such as brain and consciousness, complexity, etc. have been discussed at many workshops. Basic information of each workshop is published at the conference web site www.xssc.ac.cn. Then people who did not attend the workshop can acquire some information from related web pages, such as a summary page of each workshop which includes all primary talks, some typical questions and discussions during the workshop together with a list of all participants.

For the sake of processing, here we consider that a group discussion is composed by a set of utterance records with a structure as

$$< topic, participant, utterance, keywordsset, time >$$

Such a record indicates a *participant* submits an *utterance* with a set of *keywords* at the *time* point along a discussing process about the *topic*. Here *topic, participant, utterance* and *time* can be directly fixed as the event happens, while *keywords set*, which are manually selected and assigned to each utterance according the meaning and context of discussion. For example, the record

<E4_CoData,
xjTang,
Augmented Analytical Exploitation of a Scientific Forum,
{conference mining, Xiangshan Science Conference},
2006-10-24>

indicates that participant *xjTang* gave a talk *Augmented Analytical Exploitation of a Scientific Forum* about the topic *E4_CoData* on October 24, 2006. *Conference mining* and *Xiangshan Science Conference* are two representative keywords of the utterance. Keywords are usually indicated clearly together with the abstract by the authors during submission to general academic conferences.

Based on such a conceptual model about group discussion, two kinds of matrices are generated.

1) Frequency matrix

Two frequency matrices F_p and F_u can be acquired. Each element of matrix F_p denotes the frequency of keyword i referred by participant j during the whole discussing process. Each element of matrix F_u denotes the frequency of keyword i referred by the utterance j as shown in Table 3.1. The keywords are articulated as attributes of participants or utterances.

Given frequency matrix F_u, dual-scaling method is employed to analyze the correspondence relations between utterances and keywords. Proposed by Nishisato, dual scaling is a multi-variant statistical method that is of similar characteristics with correspondence analysis and exploratory factor analysis [2]. In Table 3.1, each element in $X = (x_1, x_2, \ldots, x_m)^T$ refers to the weight of the corresponding keyword while vector $Y = (y_1, y_2, \ldots, y_n)^T$ refers to the sum of weighted scores about the corresponding utterance. With the principal components for given relations between keywords and utterances acquired by dual scaling, both the utterances and keywords can be mapped into 2-dimensional space. As a result, a pair of utterances with more common keywords may locate closer in the 2-dimention space. Such a process may also apply to spatially mapping with relations between participants and keywords set with frequency matrix F_p.

In the computerized support tool, group argumentation environment (GAE), for group divergent work, an electronic brainstorming room (BAR) is a basic module for

Table 3.1. Frequency Matrix: Utterance sets and keyword sets

X \ Y		keyword$_1$	keyword$_2$		keyword$_m$	
		x_1	x_2	$\cdots$	x_m	
utterance$_1$	y_1	a_{11}	a_{12}	$\cdots$	a_{1i}	$y_1 = \sum_{i=1}^{m} a_{1i}x_i$
utterance$_2$	y_2	a_{21}	a_{22}	$\cdots$	a_{2i}	$y_2 = \sum_{i=1}^{m} a_{2i}x_i$
$\vdots$		$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
utterance$_n$	y_n	a_{n1}	a_{n2}	$\cdots$	a_{ni}	$y_n = \sum_{i=1}^{m} a_{ni}x_i$

(a_{ji} denotes the frequency keyword$_i$ appeared in the utterance$_j$.)

diverse idea publishing and serves as the virtual space for participants for communication and information sharing, similar to a general BBS. Besides general functions as BBS, GAE-BAR provides a visualized area to exhibit the dynamic process of discussing for one topic using the results of dual scaling method for both frequency matrices. The results for matrix F_p processed by dual scaling method are displayed as a common view; while the results for matrix F_u are displayed as a personal view. Both views serve as a visualized shared memory space where displayed the global structure of participants' joint thought about the concerned topic. Moreover, with selected participants or selected utterances, a visualized group thinking map cab be accessed via a retrospective view. Retrospect analysis in GAE-BAR mainly help users to "drill down" into the discussing process with visualized snapshots about pieces of discussion such as selected intervals of discussion or selected participants, and detect the micro-community forming, which may be useful in understanding about participants' thinking structure, story-telling or group thinking context awareness, or case studies for other problem solving.

Furthermore, based on the spatial relations in common view or personal view, clustering of keywords or utterances can be done. In GAE-BAR, centroid-based K-means method is applied to keywords clustering, while KJ method is applied to utterances grouping.

2) Matrix of Agreement or Discrepancy

The second kind of matrix concerns relationships between participants. If there are n participants at one period of time in the course, let $U_i$is the keywords set referred by the participant i, $i = (1, 2, \ldots n)$, then two matrices are acquired:

- Matrix of agreement or similarity, denoted as A_1 where $a^1_{ij} = |U_i \cap U_j|$. The element is number of the keywords shared between participants $i$and j. Obviously, $a^1_{ii} = |U_i|$.
- Matrix of dissimilarity or discrepancy, denoted as A_2 where $a^2_{ij} = |(U_i \cap \bar{U}_j) \cup (\bar{U}_i \cap U_j)|$. The element is the number of different keywords between two participants $i$and j. Obviously, $a^2_{ii} = 0$.

Both matrices are symmetrical. According to the eigenvector corresponding with the maximum eigenvalue for each matrix, we get rank of participants which may reflect participants' contributions in the course. It is estimated that eigenvector of Matrix A_1 reflects who holds more common concerns in the course, and that of Matrix A_2 reflects who is of more diverse perspectives than others. If we pay more attention to the extent of consensus or agreement, characteristics of Matrix A_1 may give some hints; if we focus on how diverse of ideas in the course, more information could be acquired from A_2. Both measures are firstly discussed in Ref. [3].

Since both matrices are non-negative, in order to ensure the existence of a unique maximum eigenvalue for each matrix, a small value, eg. 0.001, may be added to each element and the original non-negative matrix is transformed into a positive matrix. Such a change can be explained from a practical view. For matrix of agreement, those participants attending the same discussion of the concerned topic may underlie that they at least share same interests in the topic. For matrix of discrepancy, each participant is different, and even a participant himself may change ideas along the discussing process, then differences always exist even to one participant along a discussion process.

Such kind of measures may be helpful to selection of appropriate experts for relevant workshop or even problem solving later. Next, an example is given to show how those analytical methods are applied to mining of XSSC.

3.3 Augmented Analyses toward Topics on Complexity of XSSC by View of Group Argumentation

As complex systems and complexity research has becoming very hot since the 21st century, here we select some workshops, whose principal topics concentrate on complexity or complex system, to study how relevant research has been noticed and undertaken by Chinese scientists. Till May of 2004, seven relevant workshops have been held as listed in Table 3.2.

Initially the script of group argumentation includes all talks given by 17 selected scientists, whose group thinking space via common viewer in GAE-BAR is as shown in Figure 3.1.

The rectangular icon refers to the user ID of participants in the discussing process, and the oval icon refers to keyword as articulated as attributes of participants. If the mouse locates at a rectangular, then all utterances given by the corresponding participant are popped up. The more shared keywords between participants, the higher mutual relevance between them, which is reflected by the distances between participants in the map. As discussion goes on, the structure of the diagram will be changing. Then as appended all talks given by those selected scientists who attended the 7th workshop held in May of 2004, a big change happened in the common view. The final common view is displayed at Figure 3.2(a) which also shows the layout of the client window of GAE-BAR. As to details of functions or framework of GAE-BAR, please refer Ref. [3, 4]. Figure 3.2(b) shows a retrospective view with 11 selected scientists who are closely located in Figure 3.2(a). Then as indicated in the

Table 3.2. Topics about Complexity

Workshop No.	Title	ConveningTime
20	Open Complex Giant System Methodology	June 20-23, 1994
29	Theoretical method and some major Scientific Problems relevant to Natural	March 29-31, 1995
68	Theory and Practice of Open Complex Giant System Methodology	January 6-9,1997
110	Cybernetics and Revolutions in Science and Technology	December 22-23,1998
112	Complexity Science	March 18-20, 1999
190	Complex Systemin Process Engineering	September 17-19, 2002
227	System, Control and Complexity Science	May 25-27, 2004

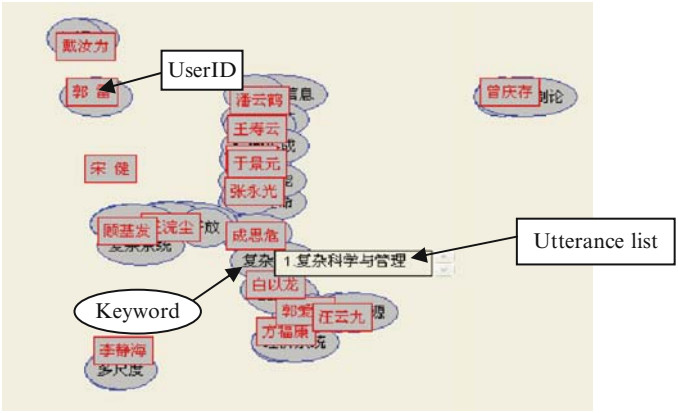

Fig 3.1. Original spatial map about "complexity" (6 workshops selected, till 2002)

middle of Figure 3.2(b), another micro community is detected where "control", "cybernetics" and "artificial intelligence" are foci. Retrospective analysis can also used by observers as a readily accessible record of the topics when facing similar issues during problem-solving process.

Next some feature functions of GAE-BAR as shown in Figure 3.2 are addressed.

(1). Recording Original Idea Provider

For all keywords proposed during a discussing process, GAE-BAR can tell users who firstly propose each keyword, when and how often that keyword is referred later, as shown in Figure 3.2(c). Such a mechanism is to check the originality of participants. In her 2-space transformation model, Boden claims the idea as P-creative if the person in whose mind the idea arises haven't had it before, no matter how many others may have had the same idea already [5]. Then the record of original idea provider is the record of individual's P-creativity among the discussing group under the same theme.

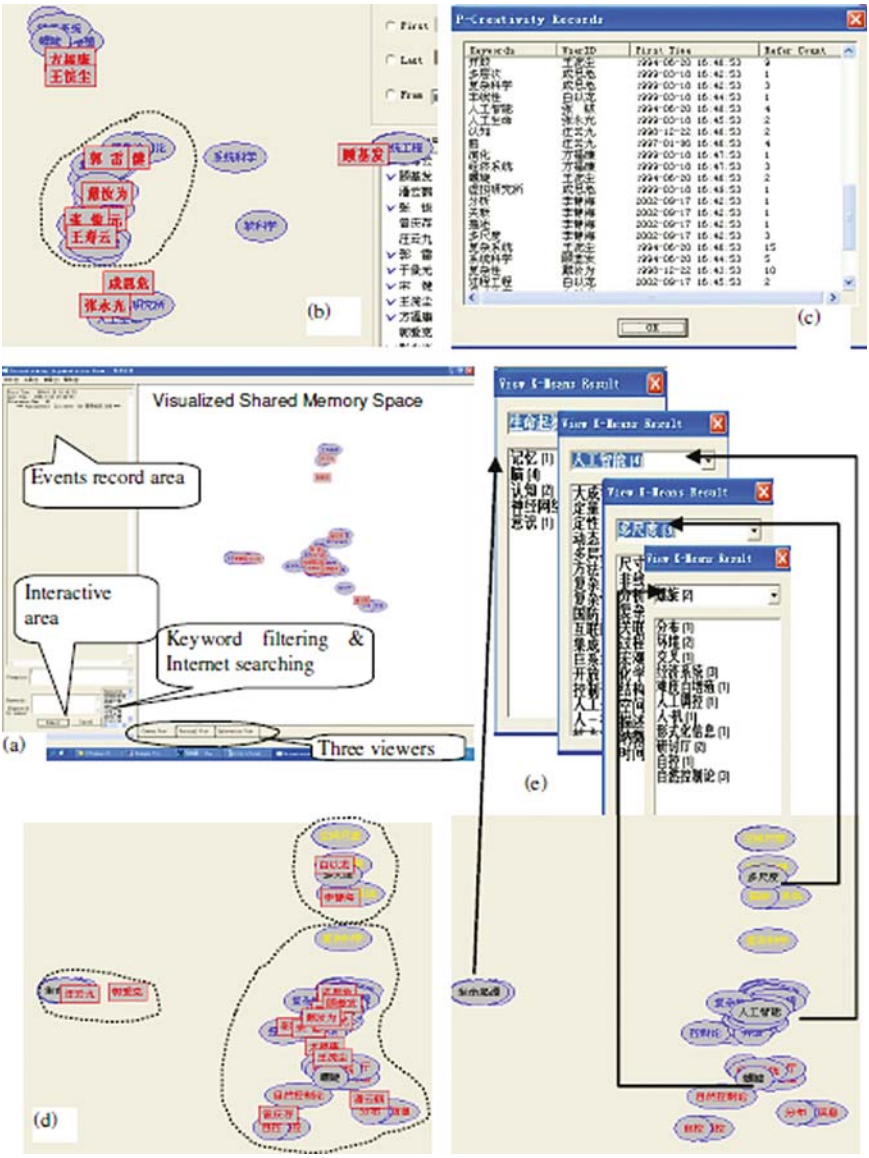

Fig 3.2. Functions of GAE-BAR
(a) Main client window; (b) Retrospective view; (c) Original keyword provider;
(d) Clustering analysis (k =4) with participants; (e) List of a cluster of keywords by (d).

The highest frequency of the referred keywords may indicate the foci of the argumentation. As shown in Figure 3.2(c), two frequently referred keywords are “complex system” (15 times) and “complexity” (9 times). Less referred keywords may also be noticed by active participants. Such a record then help users aware of those

which have not been much noticed so far (P-creativity) during the argumentation process and may change current attention with another interesting idea for further thinking. For divergent thinking, more ideas as scattered across either the common view or the personal view at GAE-BAR are always encouraged. The P-creativity recorder is then a supplement to help users aware of diverse ideas, instead only focusing those highlighted ones.

(2). Clustering of Utterances/Keywords for Perspective Formatting

Given the records of discussion process, both organizers and participants prefer somewhat summarization which is helpful to refine divergent thinking results into something that makes sense and can be dealt with more easily by human experts. Now two ways are available in GAE-BAR. One is automatic affinity diagramming (usually referred as KJ method because of its inventor, Kawakita Jiro) which maps the whole utterance set into 16×16 cells according to their spatial relationship at the persona view with a 2-dimensional structure. Those utterances which fall into same cell are regarded as one cluster. Human experts may assign one label to each cluster themselves by their judgment and discover meaningful groups of ideas from a raw list.

Another way uses k-means clustering method (k is an assumed number of clusters). Each cluster has a centroid. The keyword which is closest to the centroid of the affiliated cluster could be regarded as label of the cluster. In our example, as $k = 4$, then 4 clusters are generated (see Fig. 3.2(d)). The keyword set of each cluster is as shown in Fig. 3.2(e) where the labels for each cluster are 生命起源 "the origin of life" (the bottom window), 人工智能 "artificial intelligence" (the second bottom window), 多尺度 "multi-scale" (the second upper window) and 螺旋 "spiral" (the upper window) respectively. Then human analysts may give conceptual terms about those clusters based on machine processing results or join some clusters for meaningful summary. For example, "complexity of life and brain" could be used as the label for Cluster "the origin of life" in general. Cluster "spiral" and Cluster "artificial intelligence" can be merged as one cluster as both stress on "complex systems" and "complexity science".

Regarding the corresponding experts with those keyword clusters, expert grouping is acquired simultaneously. Moreover, such kind of correspondence provides some hints about the perspectives and even roles of those related scientists in XSSC. For example, Cluster "multi-scale" includes keywords such as spatial scale, temporal scale, micro scale, etc. and could somewhat reflect the knowledge specialty or research perspectives of both experts 白以龙 and 李静海 as shown in the left of Figure 3.2(d). This point is in accord with the reality. Similarly, Cluster "the origin of life" may indicate knowledge specialty of both expert 汪云九 and 郭雷 who are active in the field of complexity of life and brain. Such an association may be very helpful together with measures of participation as addressed next.

(3). Measures of participants' contributions to the discussions

Table 3.3 lists the evaluation of 17 participants' involvement to the discussion of the topic of complexity based on agreement and discrepancy matrices.

It is shown that the participant 戴汝为 holds highest rank at both agreement and disagreement measures, which may be justified by his active role as one of chairpersons

Table 3.3. Measures of 17 experts' participation to complexity related workshops at XSSC

The eigenvector of maximum eigen value of agreement matrix:	(0.0952, 0.5795, 0.0979, 0.3114, 0.0363, 0.0712, 0.1447, 0.2049, 0.3348, 0.4596, 0.0964, 0.0412, 0.0446, 0.3449, 0.0896, 0.0408, 0.1196)
Rank of the top five participants:	戴汝为 > 于景元 > 张钹 > 宋健 > 成思危
Meaning of the indicator:	Expert with higher rank may hold more common concerns during the brainstorming session
The eigenvector of maximum eigenvalue of discrepancy matrix:	(0.3075, 0.3288,0.2724, 0.2888,0.1848,0.1882, 0.1905, 0.2184, 0.2381, 0.2594, 0.1630, 0.2679, 0.2347, 0.2229, 0.2464, 0.2557, 0.1813)
Rank of the top five participants:	戴汝为 > 手静庵 > 成思危 > 白以龙 > 汪云九
Meaning of the indicator:	Expert with higher rank may be of more diverse perspectives during the brainstorming session

or plenary speech contributors among those 7 workshops, which furthermore exposes his big influence in complex system field in China.

Above indicators of agreement or discrepancy are just one kind of measures. Higher agreement for one participant may indicate he share more foci than those of lower ranks. If some participants always follow or response other ideas, his agreement rank may also be higher. At that moment it is better to check original idea provider (P-creativity recorder) to exclude those followers. The topics and whole discussing process should be considered when undertaking practical analyses.

3.4 Augmented Information Support for XSSC

Users of GAE-BAR can acquire external information along a discussing process via any searching engines (such as Google, Baidu, etc.) based on keywords. To help participants' concentration on discussion, an augmented information support (AIS) tool specifically designed for XSSC is implemented based on Web content mining technologies. AIS-GAE includes four functional modules, web crawler, indexing, automatic summarization and user interface for searching, to implement general Web mining tasks. Web crawler collects the web documents given the seed websites and a defined exploring depth. Different kinds of useful information, such as workshop report, etc. are extracted from those web documents according to their structure characteristics and stored into a database with the index for each page. User interface is the entry for people to search the relevant information. Figure 3.3 shows that the 137 Web pages are found and listed below by order of relevance in searching keyword "complex". Each item includes the original URL, the sentence with the highest

Fig 3.3. Main page of user interface of AIS-GAE

relevance rank at the corresponding page referred by URL, the link to abstract of that page, the link to participants list and the link to related conferences.

The result lists really expose a rough scenario that how Chinese scientists approach complexity or complex system from different disciplines and how those scientists interact across different disciplines via the platform provided by XSSC. For example, the theme of workshop indicated by the first item in Figure 3.3 is about system and control, Item 2 is about complexity in the brain, Item 3 is from the perspective of complex system modeling and system engineering practice, and Item 4 is about a workshop on medical sciences and life, etc.

According to the specific structure of XSSC web page, a list of participants is extracted from each summary page of one workshop. Then a full list of all participants is acquired. Therefore besides an abstract of related introduction or overview of the workshop, a list of participants together with their affiliations is also provided by AIS-GAE. Figure 3.4 displays the summarization of the original web document listed as the 3rd item of all 137 results as shown in Figure 3.3. Figure 3.5 shows part participants of that workshop. Figure 3.6 provides a list of other XSSC workshops at the same cluster as the referred workshop. Clustering of XSSC Web texts is discussed at Chapter 9 in this book.

AIS-GAE could provide help for three kinds of people relevant to XSSC, i) the workshop organizers who can search past relevant workshops information and

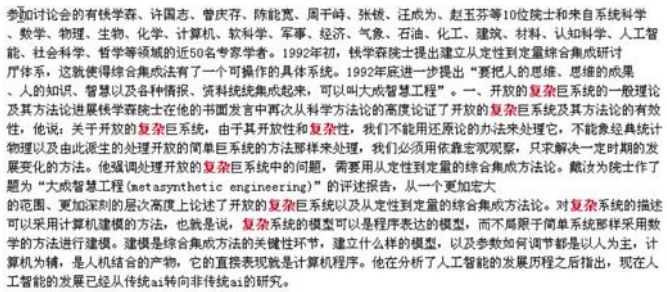
参加讨论会的有钱学森、许国志、曾庆存、陈能宽、周干峙、张钹、汪成为、赵玉芬等10位院士和来自系统科学
、数学、物理、生物、化学、计算机、软科学、军事、经济、气象、石油、化工、建筑、材料、认知科学、人工智
能、社会科学、哲学等领域的近50名专家学者。1992年初，钱学森院士提出建立从定性到定量综合集成研讨
厅体系，这就使得综合集成法有了一个可操作的具体系统。1992年底进一步提出“要把人的思维、思维的成果
、人的知识、智慧以及各种情报、资料统统集成起来，可以叫大成智慧工程”。一、开放的复杂巨系统的一般理论
及其方法论进展钱学森院士在他的书面发言中再次从科学方法论的高度论证了开放的复杂巨系统及其方法论的有效
性，他说：关于开放的复杂巨系统，由于其开放性和复杂性，我们不能用还原论的办法来处理它，不能象经典统计
物理以及由此派生的处理开放的简单巨系统的方法那样来处理，我们必须用依靠宏观观察，只求解决一定时期的发
展变化的方法。他强调处理开放的复杂巨系统中的问题，需要用从定性到定量的综合集成方法论。戴汝为院士作了
题为“大成智慧工程(metasynthetic engineering)”的评述报告，从一个更加宏大
的范围、更加深刻的层次高度上论述了开放的复杂巨系统以及从定性到定量的综合集成方法论。对复杂系统的描述
可以采用计算机建模的方法，也就是说，复杂系统的模型可以是程序表达的模型，而不局限于简单系统那样采用数
学的方法进行建模。建模是综合集成方法的关键性环节，建立什么样的模型，以及参数如何调节都是以人为主，计
算机为辅，是人机结合的产物，它的直接表现就是计算机程序。他在分析了人工智能的发展历程之后指出，现在人
工智能的发展已经从传统ai转向非传统ai的研究。

Fig 3.4. Abstract of the original Web page with the searching keyword “complex” highlighted in red

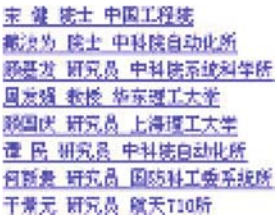
宋 健 院士 中国工程院
戴汝为 院士 中科院自动化所
顾基发 研究员 中科院系统科学所
国家强 教授 华东理工大学
顾国庆 研究员 上海理工大学
谭 民 研究员 中科院自动化所
何新贵 研究员 国防科工委系统所
于景元 研究员 航天710所

Fig 3.5. Participant list of a XSSC workshop

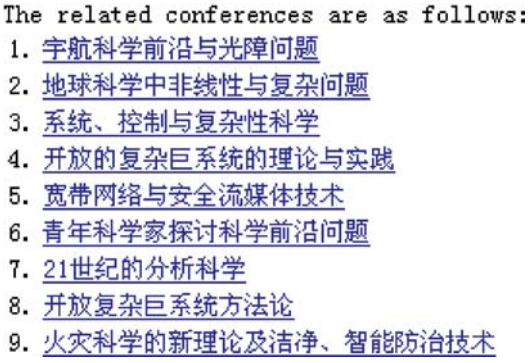
The related conferences are as follows:
1. 宇航科学前沿与光障问题
2. 地球科学中非线性与复杂问题
3. 系统、控制与复杂性科学
4. 开放的复杂巨系统的理论与实践
5. 宽带网络与安全流媒体技术
6. 青年科学家探讨科学前沿问题
7. 21世纪的分析科学
8. 开放复杂巨系统方法论
9. 火灾科学的新理论及洁净、智能防治技术

Fig 3.6. Related workshops

acquire a whole vision toward XSSC; ii) the reviewers of a workshop application who can compare with past workshops regarding available similar themes and give their judgments about necessity and originality of applications; and iii) the invited participants of a workshop who may prepare their talks with more visions and engage actively during the workshop.

As time goes on, it is natural that the changes in position and even organization will happen to some participants. Such kind of additional information about participants will be pushed to the users as they try to search a participant using AIS-GAE. For example, we type a name such as 于景元 instead of a general keyword and start

Fig 3.7. Search a participant by AIS-GAE

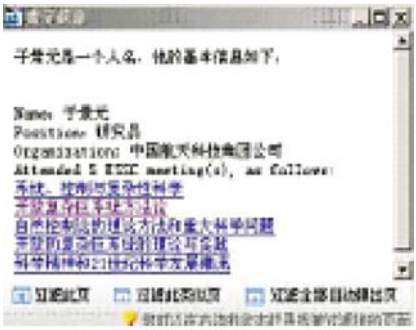

Fig 3.8. Push information by AIS-GAE

searching. The searching result is as shown in Fig. 3.7. Here 6 relevant Web pages are found.

Moreover, as the inputted word is detected as a name based on participants list, AIS-GAE then pushes relevant information of that participant to the users. Due to popup blocks, push information may need to be released (see a button of release in Fig. 3.7). As shown in Fig. 3.8, the popup information includes the latest position and affiliation of the relevant participant together with all XSSC workshops he has attended (here, Professor 于景元 had participated 5 workshops).

Most participants of general XSSC workshops are nominated by workshop presidents and then invited by the XSSC organizers. In spite of a tenet of open-mindedness and cross disciplines, there still exists a tendency to recommend popular scientists with higher position, better reputation or from famous organizations among many XSSC workshops. Based on participant list, a human network can then be constructed by AIS-GAE. In the network, the vertex refers to a participant. If two

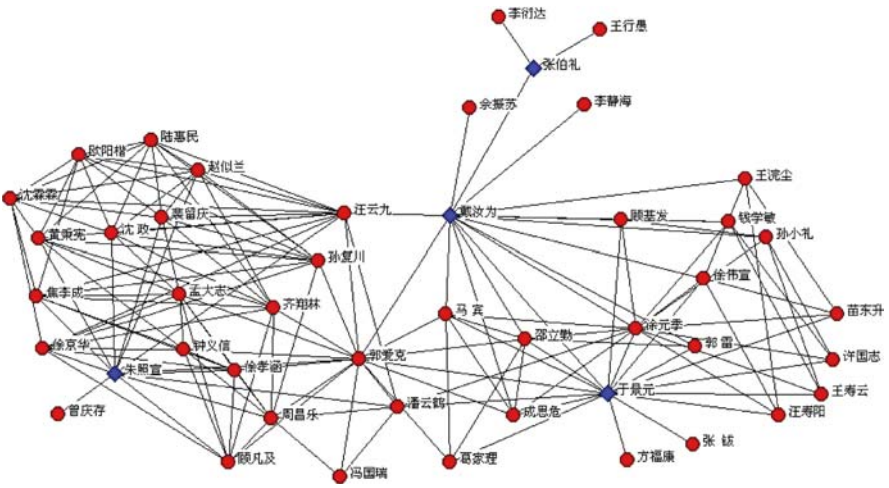

Fig 3.9. Human network composed by participants who simultaneously attend at least twice workshops (: cutpoints)

participants i and j attend a workshop simultaneously, then there is an edge between vertex i and vertex j. The strength of the edge is the frequency of co-occurrence of both participants along all concerned workshops. Fig. 3.9 shows a human activity network where the strength of the edge is at least greater than 2, which means both participants sharing one edge had been simultaneously attended those complexity workshops at least twice.

Here social network analysis is naturally applied to detect some features from that network. For example, 4 cutpoints are found in Fig. 3.9. Those corresponding scientists are 戴汝为 (middle), 于景元 (right), 张伯礼 (upper) and 朱照宣 (left), which reflects their important roles in the complexity related workshops, such as the former two scientists had been served as presidents for several times. In consideration of their academic background, the cutpoint scientists may also be regarded as gatekeeper of different school about complexity research in China. If combined with those measures of participations and participants group provided by GAE-BAR, more information about participants may be acquired.

3.5 Concluding Remarks

In this paper, we focus on mining information from a knowledge repository, Xiangshan Science Conference, a famous scientific forum on frontiers of science and technology in China. Unlike traditional simple statistics and qualitative evaluation toward conference, here adopted some augmented analytical methods, such as

- Visualized thinking space or structure of a group of scientists toward specified topics;

- Various clustering about the workshop discussion for effective summarization;
- Measuring the contributions of participants, which aims to provide more help to organizers in inviting or selecting participants;
- Augmented information support by Web content mining, especially push information about participants;
- Participants network analysis to detect influential scientists or find gatekeepers about some research streams.

Those versatile methods have been integrated into a computerized tool, group argumentation environment (GAE), which aims to support divergent group thinking process for the emergence of a *ba* for knowledge creation. Here, GAE-BAR and AIS-GAE may serve as a conference assistant system to help organizers and researchers aware of those ignored information, especially the productive process of series workshops where burgeoning new disciplines and scientists. Not withstanding complaints never fade about the low efficiency of group meetings in reality, they are still feasible and effective ways for communication and information sharing, opinion collection and expert knowledge acquisition. Therefore computerized support for group work, especially toward the awareness of group working context, are of continuous research.

With GAE, in-depth studies toward more facets about XSSC may help to expose more precious information from such a scientific think tank, and then serve as reasons to maintain the scale of those dialogues and facilitate policy making toward fundamental research in China.

As a matter of course, GAE may be of advantages in supporting small-scale group argumentation at current stage. With practice of GAE, especially on XSSC, there emerged new issues which are of in-depth research. For example, about clustering of participants' ideas, similar with constructing human activity network, a keyword network can also be constructed where the vertex refers to a keyword, and if both keyword i and keyword j occur simultaneously in one utterance, then an edge exists between two vertexes while the weight of the edge refers to the frequency of co-occurrence of both keywords along all the concerned workshops [6]. With the community (subgroup) of keywords detected from this network, more perspectives toward those concerned topics may be acquired together with clustering of keywords taken in GAE-BAR. Moreover, analysis of structure about those communities (keywords or participants) detected from either network can be studied, for example using relevant algorithms proposed by Newman and Girvan [7].

It is worth indicating that all analyses toward XSSC introduced here were undertaken after the finish of those workshops. It is our motivation to apply GAE to those on-going workshops which may be especially better to facilitate the construction of a context of show-and-tell for both participants and other interested people with things to "show" who want to "hear" tell, and for people who give tell to things shown. Augmented support may facilitate active interactions between participants by think-and-play and enable productive work for both scientists and organizers. Then more diverse and visualized association techniques about the discussion topics will be under further exploration or adopted and integrated into GAE. For example,

ideas and technologies of chance discovery proposed by Japanese scientists [8] to detect possible networks to show "islands and bridges" (hidden chance) is worth some synthesizing with current GAE conferencing mining methods. Whatever's about next research, our original idea is to explore more structuring approaches toward unstructured problem solving.

Acknowledgments

The authors are grateful to faculty members of Xiangshan Science Conference Office for their encouragement to this study, which is supported by Natural Sciences Foundation of China under Grant No. 70571078 and 70221001.

References

1. Thagard, P.: How Scientists Explain Disease, Princeton University Press, Princeton, New Jersey (1999)
2. Nishisato, S.: Analysis of Categorical Datadual scaling and its applications. University of Toronto Press, Toronto, (1980) 1-53
3. Tang, X. J., Liu, Y. J.: Exploring computerized support for group argumentation for idea generation. In: Nakamori, Y. et al. (eds.): Proceedings of the 5th International Symposium on Knowledge and Systems Sciences, Japan (2004) 296-302
4. Tang, X. J., Liu, Y. J.: Computerized Support for Idea Generation during Knowledge Creating Process. In Cao, C. G., Sui, Y. F. (eds.): Knowledge Economy Meets Science and Technology (KEST'2004), Tsinghua University Press, Beijing, (2004) 81-88
5. Stefik, M., Smoliar, S.: The Creative Mind: Myths and Mechanisms: six reviews and a response, Artificial Intelligence, Vol.79, No. 1 (1995) 65-67
6. Tang, X. J., Liu, Y. J.: Computerized Support for Qualitative Meta-synthesis as Perspective Development for Complex Problem Solving. In Adam, F. et al. (eds.): Creativity and Innovation in Decision Making and Decision Support (proceedings of IFIP WG 8.3 International Conference on Creativity and Innovation in Decision Making and Decision Support), Vol.1, Decision Support Press, London, (2006) 432-448
7. Newman, M E J, Girvan, M.: Finding and Evaluating Community Structure in Networks, Physical Review E (2004) 69:026113.
8. Ohsawa, Y., Tsumoto, S. (eds) Chance Discoveries in Real World Decision Making, Springer-Verlag, Berlin (2006)