ISSN 2320 - 2602
Volume 2, No.8, August 2013
International Journal of Advances in Computer Science and Technology
Available Online at http://warse.org/pdfs/2013/ijacst03282013.pdf

A GRAPH-BASED INTERACTION PATTERN DISCOVERY FOR HUMAN MEETINGS

D. Karthika[1], R. RangaRaj[2]
[1]Research Scholar,Department of Computer Science, Hindusthan College of Arts and Science, Coimbatore, India. d.karthi666@gmail.com
[2] Head, Department of Computer Science, Hindusthan College of Arts and Science, Coimbatore,India. rraj75@rediffmail.com

ABSTRACT

Mining Human Interaction flow in meetings or general representation of any interaction face to face to meetings is useful to identify the person reaction in dissimilar situation. Activities represent the natural history of the individual and mining methods help to analyze how person delivers their opinion in different ways. Meeting interactions are categorized as propose, comment, acknowledgement, request-information, ask-opinion, positive -opinion and negative opinion. From this Detecting semantic knowledge is significant. Existing system data mining technique to detect and analyze frequent interaction patterns to discover various types of new knowledge on interactions. An interaction flow in user is represented as tree. Tree based pattern mining algorithms was planned to analyze tree structures and extract interaction flow patterns. This work has extend an interaction tree based mining algorithm in two ways: Human interaction flow in a session extraction of the similar events with temporal data mining techniques, it extract the temporal patterns from the captured substance of time series of dissimilar meetings in specific period of time. After that a graph based mining method is proposed to extract the frequent patterns and mining the best meeting pattern. Graph-based Substructure pattern mining which discovers frequent substructures patterns from the face to face meeting not including applicant invention. It builds a new lexicographic order among graphs or tree representation and maps each graph to a unique minimum DFS code as its canonical label with human interaction pattern representation. Based on this lexicographic order adopts the depth-first search approach to extract frequent connected subgraphs proficiently. An experimental result shows that proposed Graph-based Substructure pattern mining algorithm substantially outperforms than the previous tree based mining algorithms.

Keywords: Human Interaction Flow, Tree based mining, Frequent itemset mining, Graph based mining and Clustering based representation.

1.INTRODUCTION

Data mining, which is an important technique for discover original information, extensively adopt in numerous fields such as bioinformatics, marketing and security. KDD is the process of discovering original patterns from large data sets concerning the methods at the grouping of artificial intelligence, machine learning, statistics and database systems [2] . The process of data mining is to extract knowledge from a dataset in a human-understandable structure. In the societal dynamics such as person interaction is the one of the most important thing for considerate, how a human's behavior or human actions under the assembly and determining whether the assembly was well organized or not is one of the main issues in meetings. To overcome these numerous methods have been anticipated to originate the interaction of flow in the meeting at each human. Individual person interaction is one of the most significant distinctiveness of group social dynamics in meetings. We are developing a face to face meeting such as when user enters the shopping in the websites and give their opinion for is such as propose an new idea, generous comments, expressing a positive opinion, negative opinion etc., that implies user part, thoughts, or purpose toward a topic. To more understand the human activities and interfering of the human interactions in meetings, have to determine higher level of semantic information such as which interactions flow often occurs in a discussion and relationship among interactions in meetings. It will help to describe imperative a pattern of interaction [1].

In this study, data mining technique is detected and examine frequent interaction patterns at meetings. Human interaction flow in a conversation meeting can be representing as a tree structure. Motivated by tree-based mining [3] [4] considered interaction tree pattern mining algorithms to evaluate tree structures and extract interaction flow patterns. The human interaction flow that appears frequently reveals relationships between different types of interactions. The results are useful for Cognitive science researchers could use them as field information for further examination of human interaction. Furthermore the exposed patterns can be utilized to assess whether a meeting discussion is well-organized and to contrast two meeting consideration using interaction flow as a key characteristic.

Additionally have expanded an Graph-based Substructure pattern mining which target to reach the best pattern .To the best of our knowledge first it is the DFS in frequent subgraph mining. It can be performed in two ways: DFS lexicographic order and minimum DFS code which structure a novel canonical labeling system to sustain DFS search. It discovers all the frequent subgraphs not including candidate generation. It combines the emergent and examination of frequent sub graphs patterns into one process, thus accelerate the mining procedure.

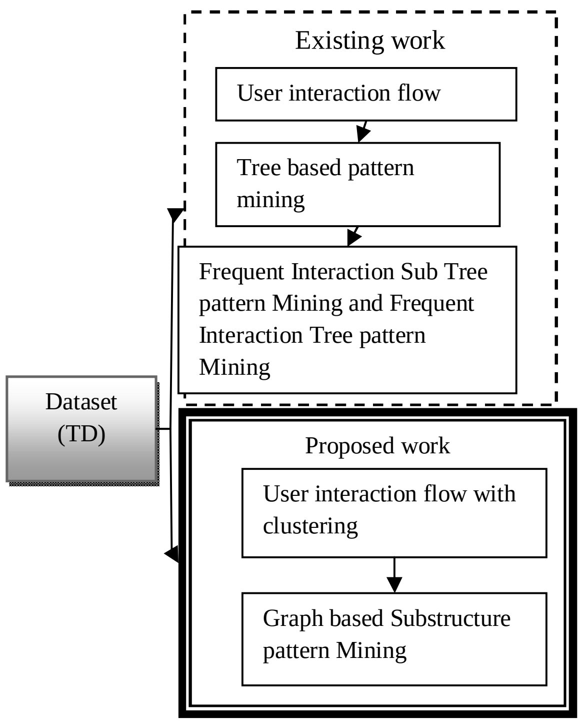


Figure 1: Architecture of the system

2. BACKGROUND STUDY

Human interaction is one of the most significant individuality of group social dynamics in meetings. Unlike corporal interactions such as, turn-taking and addressing the human communications are incorporated with semantics. Adopt a collaborative approach for capturing interactions [5] mostly focus on detecting physical interactions among participants without any relationships with topics. Therefore they cannot obviously conclude participant's attitude. Every increasing amount of recorded gathering data is motivating the need for the completion of tools to proficiently access and quickly retrieves significant pieces.

To discover the frequent patterns in a tree H. Aoyama [1] introduces a novel algorithm to discover all frequent patterns subtrees in a tree-plant with a novel data structure called scope-list. In this scheme TREEMINER through a pattern matching tree mining algorithm were contrasted by H. Aoyama. It also present a relevance of tree mining to examine real web logs for usage patterns. In earlier work XSpanner [6] systematically expand the two algorithm pattern growth methods for drawing out frequent tree patterns. Experimental results show that the recently developed algorithm out-performs TreeMinerV. Furthermore XSpanner is considerably faster for frequent tree pattern mining and scalable while the patterns are not too difficult. Even though there are numerous complex patterns in the data set.

Modeling and tracking a person's focus of attention is helpful for numerous applications. For multimodal person processor interaction, the user's concentration of attention can be used to make a decision his/her message purpose. We address the difficulty of tracking the illustration focal point of attention to participants in a meeting; a key aspect [8] of interactive meetings is the present speaker, who is, by definition, changing frequently. Follow a line of investigation indicates with the purpose of viewing recorded meetings, the mixture of an automatic speaker and fixed context views does not provide sufficient information. Theoretical out two high-level requirements for gathering viewing interfaces which are not fully addressed by present meeting viewing system:

Requirement 1: The interface is supposed to properly express to the user who an attendee is looking at.

Requirement 2: The context view is supposed to permit the user to focal point on any attendee or any part of the meeting room.

A smart meeting system which record also analyze the generated audio designed for future viewing, the directly above mentioned topic develops a great challenge in recent years. An effective smart meeting system relies on different technologies, extending from devices as well as algorithms. This system [5] presents a study of existing research as well as technologies, including smart meeting system, meeting capture ,semantic processing, meeting recognition, and evaluation methods. This article also refer to various issues of all possible ways towards extend the capabilities of current smart meeting systems.

An omni directional camera [4] is used to capture the scene all over the place a meeting table. Here actual time face tracker is used to detect as well as track participants in the panoramic image. Moreover, neural networks (NN) are used to compute head posture of every one person simultaneously on or after the panoramic image. At that time use a Bayesian approach to estimate a person's focus of attention from the figured head pose. Since Hand-recorded notes have several drawbacks. Taking notes is time consuming; that one requires extra focus also therefore reduces one's attention. For this purpose as well as remarks tends to exist unfinished then moderately summarized.

Layer models [8] characteristic the actions of persons in meetings with supervised HMM learning and low-level audio-visual features. Numeral of options that clearly model with positive aspects of the data. HMM model subsequent layer the collection actions with unsupervised learning. These two layers are connected by a set of probability-based features fashioned by the individual action. The methodology was assessed on a set of multimodal turn-taking group actions, using a public -hour meeting corpus. From the outcome says those layered frameworks are compared to a variety of baseline methods.

The head gestures [3] including a Wavelet-based technique beginning magnetic sensor signals. The straightforward utterance of a few platitudes is detected with data captured by lapel microphones. Experiments were conduct on four-person conversations; it validates the effectiveness of the structure in discovering interactions such as question-and-answer and addressing behavior by back-channel response. Face-to-face conversation is one of the most basic form of message in our time and is used for conveying/sharing information and making decisions. To enhance our announcement ability beyond conversation on the spot, the automatic analysis of discussion scenes is a basic technical to realize announcement via social agents and robots.

A pattern mining method [9] to mine significant patterns of communication from a data set that contain primitive in sequence of interaction like gazing or utterance. The technique extracts accidental patterns of communication fashioned by a set of primitive events while such a pattern occur further than randomly an interaction corpus is used for numerous purposes. With machine-readable indices, we can recapitulate a set of events and look for scrupulous events because they include a variety of kinds of context information. It is used in numerous domains, e.g., medical, genetic research and marketing and has formed many good results. The patterns of interaction fall into two category. One is a concurrently occurring pattern. This pattern can be representing by a mixture of events. Even in such a complex case, though the structure of the pattern cascade into individual of the two category types.

The Research [10] proposed a rigorous arrangement of spoken language, which is now expressed in grammars, remain mainly unnoticed all the way through thousands and thousands of years. It seems as a result quite probable that to a great extent structure in human interactions remains undiscovered. Since appreciative will have to deal with difficult real-time stream of behavior alongside fashioned by two or more persons. It is unspecified that interaction imply additional than one party somewhere, minimally, one is influencing one another. The purpose and place of an individual inside the interaction system is affected by its interactions

3. HUMAN INTERACTION FLOW

Discovering semantic information is important for sympathetic and interpret how people can act together in a meeting discussion. Capturing all of this informal assembly information has been a topic of research in several communities over the past decade. In this paper describe the human interaction flow and discovering frequent patterns of human interaction in meeting discussions at same meeting by using graph based substructure frequent. The removal results would be useful for summarization, indexing, and comparison of meeting records.

3.1 Human Interation

The description of human interaction categorizes obviously vary according to the practice of the meeting or the type of the meetings. In this investigation mainly focus on the face to face meetings that are online purchase through ecommerce interactions. For a numerical structure for conceptualizing, investigate and scheming interaction flow, generate a set of interaction types based on category scheme: suggest, remark, acknowledgement and requestInfo, askOpinion, posOpinion and negOpinion.

3.2 Human Interation Flow

Human interaction flow in face to face meeting is premeditated as tree .An interaction flow is a list of all interactions in a conversation with the relationship between them. An interaction flow is a list of all interactions in a conversation session with triggering relationship between them. L = {SUG; REM; ACK; REQ; ASK; POS; NEG} Labels are abbreviated names of interactions, i.e., SUG-suggest, REM—Remarks, ACK—acknowledgement, REQ—request Info, ASK—askOpinion, POS—posOpinion and NEG-negOpinion .Examples of interaction trees shown in the figure 2

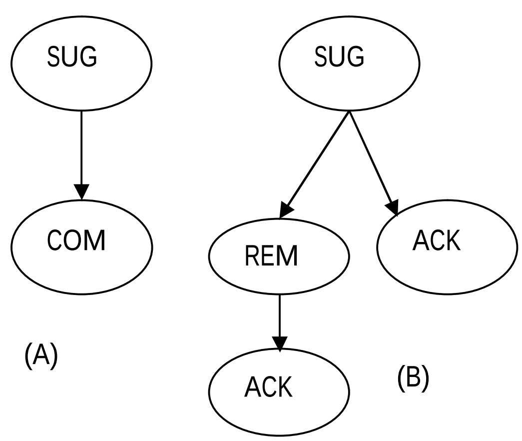


Figure 2: Examples of tree representation human interaction flow

3.3 Tree Based Pattern Mining Algorithm

Human interactions such as propose an idea, generous remarks and expressing a positive opinion, point out user intention in the direction of a theme or role in a discussion. Human interaction flow in a conversation session is represented as a tree. Tree-based relations mining algorithms are considered to examine the structures of the trees and to extort interaction flow patterns. Designed a tree based pattern mining algorithm for interaction flow mining. It formulate the frequent tree pattern mining algorithm for every node in the tree .For every tree in TD the algorithm first interactions the places of siblings to produce the full set of isomorphic trees (ITD). The principle of generating isomorphic trees is to easiness string matching. Following generate the isomorphic trees then calculates support values of alltrees at Steps 2-3. In Step 4, it selects the trees whose supports are larger than σ and detect isomorphic trees inside them. If m trees are isomorphic, it selects individual of them and rejects the others. It finally outputs all frequent tree patterns with respect to σ .

Algorithm 1: FITM (TD, σ) (Frequent interaction tree pattern mining)

Input: Tree database (TD) and a support threshold σ

Output: Frequent tree configurations with respect to σ

Procedure:

(1) Scan database TD and generate its full set of isomorphic trees (ITD)

(2) Scan database ITD and count the numeral of occurrences for each tree t

(3) Calculate the support of each one tree

(4) Select the trees whose supports are higher than σalso detect isomorphic trees; if m trees are isomorphic, select one of them and discard the others.

(5) Output the frequent trees

Where, TD: A dataset of interaction trees in the flow, ITD: The full set of isomorphic trees to TD

T:Tree, t^k: The subtree with k nodes, i.e K-subtree , C^k: Candidate set with k-nodes.
F^k: Frequent set with k-subtreesσ- support threshold minsup

3.4 Frequent Interactiion Subtree Pattern Mining

It primary calculates the support of every node and selects the nodes whose supports are larger than σ to form the set of frequent nodes, $F^1$from Steps 2-3. It then adds a frequent node to accessible frequent i-subtrees to make the set of candidates with i + 1 node at Steps 4-8. If there are any trees whose supports are larger than , it selects them to form F^{i+1} and repeats the procedure from Step 4, otherwise it stops to output of frequent subtrees. In Step 7we join t^i and t^1 to generate the candidate subtree set of size.

Algorithm 2: FISTM (TD) (Frequent interaction subtree pattern mining)

Input: Tree database (TD) and a support threshold σ

Output: Frequent subtree configurations with respect to σ

Procedure:

(1) i<-0

(2) Scan database TD, calculate the support of each node

(3) Select the nodes whose chains are longer than σ to form F^1

(4) i<- i +1

(5) For each tree t^i in F^i, do

(6) For each node t^1 in F^1, do

(7) Join t^i and t^1 to generate C

(8) Subtree Support Calculating (TD; t^{i+1})//calculate the support of each tree in C^{i+1}

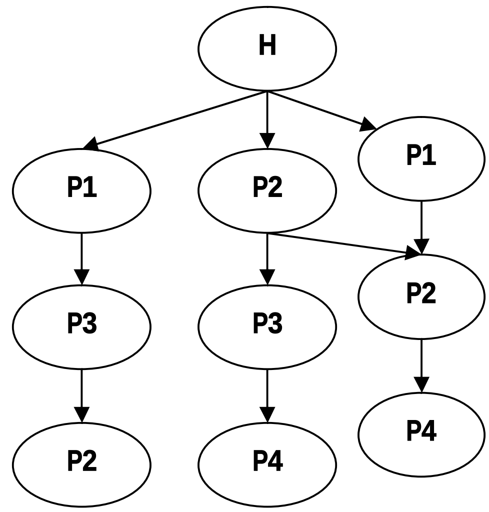
(9) if there are any trees whose supports are larger than ,then select them to form F^{i+1} and return to Step (4)

(10) Else output the recurrent subtrees whose supports are higher than σ

3.5 Graph-based Substructure pattern mining and interaction capturing with clustering

Extraction the suitable events of flow patterns form the face to face to representation from individual points we represent the user interaction flow in hierarchical representation with top down approach by clustering events having spatial and temporal relationships. Where H denoted the head node denotes Root and P denotes the person who participate in the face to face meetings. Participating persons are number from left to right, origin denoting the person who organizes the communication. Figure 3 specifies Head, person 4 and Person 2 are initiating the new statement with propose. Tree Hierarchy represents the flow in which the person represents their comments.

In this section introduce a method gSpan; with mapping each graph to a DFS code sequence, building an original lexicographic ordering amongst these codes and constructing a search tree based on this lexicographic order. Before formulation of the graph based structure tree based mining have first study the basic concepts of the depth first search and then proceed the proposed system.



.

Figure 3: Specifies Head, person 4 and Person 2

Depth first search (DFS)

Given a DFS tree T for a graph G, an edge sequence e_i can be constructed based on e_i based on t such that $e_i <_t e_{i+1}$ where i=0,1…|E-1| and e_i is called as depth first code (DFS) and it is denoted as (G,T).

DFS Lexicographic Order is a linear order defined as. If $\propto = \text{code}(G_\propto, T_\propto) = (a_0, a_1, \ldots, a_m)$ and $\propto = \text{code}(G_\beta, T_\beta) = (b_0, b_1, \ldots, b_n)$ where $\propto, \beta \in Z$ then $\propto \leq \beta$ iff either of the following true

- $\exists t, 0 \leq t \leq \min(m, n), a_k = b_k, f\ or\ k < t\ a_t < e\ b_t$
- $a_k = b_k, f\ or\ 0 \leq k \leq m, and\ n \geq m$

Given label set L should contain the tree or user based patterns from the meeting it may be opinion and request, reply of the user and it include a finite numeral of graphs or patterns of the users. Because we only judge frequent subgraphs in a finite user interaction flow in the transaction database (TD), the range of a DFS Code Tree is finite. DFS Code Tree, the n_{th} level nodes include DFS codes of (n-1) edge graphs. All the way through depth- first search of the code tree, all the minimum DFS codes of frequent subgraphs of the user interaction patterns can be exposed. That is, all the frequent subgraphs of the user interaction patterns can be discovered in this manner. It can be shown in figure 3 the darken nodes include the same graph but different DFS interaction flow patterns and then s' is not the minimum patterns of the user.

Graph based sub structure based pattern mining uses a sparse adjacency list representation to store graphs. Algorithm 3 outlines the steps of the framework which is self-illuminating. Presuppose that contain a label set {A,B,C,…….. } for vertices and {a,b,c,….}for edges. The algorithm shown below first discover the user frequent patterns from the interaction flow of the user or individual person containing an edge $A \overset{\alpha}{\rightrightarrows} A$ and Then discover the subsequent patterns containing $A \overset{\alpha}{\rightrightarrows} B$, but not any $A \overset{\alpha}{\rightrightarrows} A$ .this repeat until all the frequent user interaction flow are discovered at the face to face meetings The transaction database (TD) Procedure algorithm shown in algorithm 3 and when the subgraph turns to be larger (Subprocedure 1-line 8, simply graphs which contain this subgraph are measured. TD_S Means the set of graph in which s is a subgraph). Subgraph Mining is simply call to produce the graphs and find all their frequent children. Subgraph mining stops thorough moreover while the support of a graph is less than minsup or its code is not a minimum code which means this graph and all its children comprise been generate and exposed earlier than.

Algorithm 3 GraphSet (TD,S).

1. Sort the labels in TD based on their frequency
2. Remove infrequent vertices and edges;
3. Relabeling the remaining vertices and edges;
4. $S' \leftarrow$ $all\ frequent\ 1 - edge\ graphs\ in$ TD
5. sort S' in DFS lexicographic order
6. $S \leftarrow S'$
7. For each edge $e \in S'$
8. initialize s with e, set s.TD by graphs which contains e;
9. call Subgraph Mining(TD,S,s)
10. TD ←TD-e
11. If TD<minsup
12. Break

Subgraph Mining(TD,S,s)

1. If s≠ $\min(s)$
2. Return
3. $S \leftarrow S \cup \{s\}$
4. Enumerate s graph in TD and count it siblings
5. For each c,c in child s' do
6. If support(c)≥ $minsup$
7. $s \leftarrow c$
8. Subgraph Mining(TD_s,S,s)

4. RESUTS AND DISCUSSION

In this section measure the accuracy of the tree based mining with frequent interaction and graph substructure pattern mining with human interaction. Based on the minimum support count value the number of frequent interaction pattern mining changes at both systems. Accuracy comparison results also changes based on number of sessions. In the figure 4 measure the accuracy of the system between Frequent interaction tree mining (FITM), Frequent interaction subtree mining (FISTM),graph substructure pattern mining (GSPM).Number of sessions are represented in x axis and their corresponding accuracy of each system are measured in y axis. The accuracy of the each system values are tabulated at Table 1.

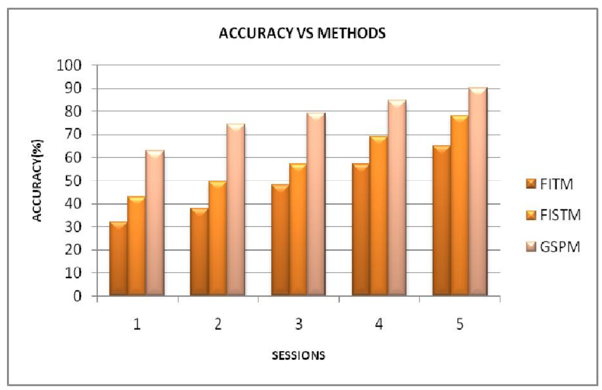


Figure 4: Accuracy comparison of FITM, FISTM and GSPM

Table 1 : Accuracy comparison of FITM, FISTM and GSPM

Number of sessions	Accuracy		
	FITM	FISTM	GSPM
1	32	43	63
2	38	49.5	74.5
3	48	57	79
4	57	69	84.5
5	65	78	90

In the figure 5 number of discovered frequent subtrees between Frequent interaction tree mining (FITM), Frequent interaction subtree mining (FISTM),graph substructure pattern mining (GSPM).Minimum threshold values are represented in x axis and their corresponding number of discovered frequent subtrees are measured in y axis. The number of discovered frequent subtrees of each system values are tabulated at Table 2.

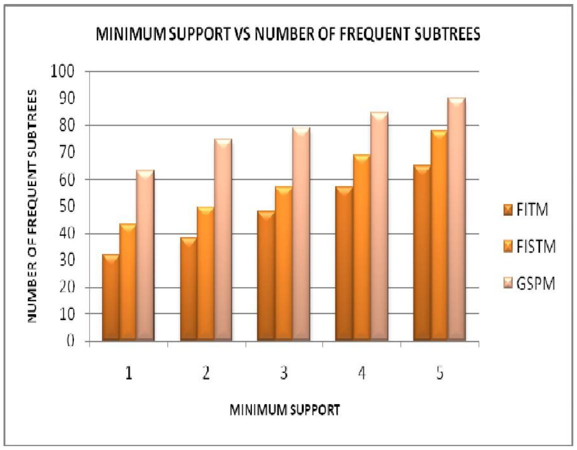


Figure 5: Support threshold and the number of discovered frequent subtrees.

Table 2 : Support threshold and the number of discovered frequent subtrees.

Minimum support value	Number of Frequent patterns		
	FITM	**FISTM**	**GSPM**
1	125	105	83
2	75	65	47
3	65	48	36
4	50	37	25
5	40	25	18

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

This paper have proposed an interaction based tree mining for human interaction flow from face to face meetings ,first cluster the user behavior using the hierarchical clustering method and then apply the graph based substructure tree based mining with lexicographic ordering system by using Depth first search (DFS) for mining frequent subgraphs in large transaction database . Performance results shows that it is capable to mine large frequent subgraphs in a bigger graphset with lower smallest amount supports than previous study.

5.2 Future Work

As future work have encompass an design to integrate more contexts like lexical cues in the discovery process in order to improve the appreciation accuracy and also improve the frequent patterns results by Appling the pruning strategies like datamining methods such as optimization, learning methods for reviewing the human interactions.

REFERENCES

[1] Z.W. Yu, Z.Y. Yu, H. Aoyama, M. Ozeki, and Y. Nakamura, “Capture, Recognition, and Visualization of Human Semantic Interactions in Meetings,” Proc. Eighth IEEE Int’l Conf. Pervasive Computing and Comm. (PerCom ’10), pp. 107-115, Mar.-Apr. 2010.

[2] Q. Yang and X. Wu, “10 Challenging Problems in Data Mining Research,” Int’l J. Information Technology and Decision Making, vol. 5, no. 4, pp. 597-604, 2006.

[3] M.J. Zaki, “Efficiently Mining Frequent Trees in a Forest: Algorithms and Applications,” IEEE Trans. Knowledge and Data Eng., vol. 17, no. 8, pp. 1021-1035, Aug. 2005.

[4] C. Wang, M. Hong, J. Pei, H. Zhou, W. Wang, and B. Shi, “Efficient Pattern-Growth Methods for Frequent Tree Pattern Mining,” Proc. Pacific- Asia Conf. Knowledge Discovery and Data Mining (PAKDD ’04), pp. 441- 451, 2004.

[5] Christian Becker and Zhiwen Yu, ‘Tree-Based Mining for Discovering Patterns of Human Interaction in Meetings,’ Proc. Eighth IEEE Int’l Conf. Knowledge Discovery and Data Mining (PAKDD’10), pp. 107-115,2012.

[6] M. Barnard, S. Bengio,D. Gatica-Perez, G. Lathoud, I. Mccowan, D. Zhang, ‘Automatic Analysis of Multimodal Group Actions in Meetings,’ IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp. 305-317,2005.

[7] P. Chou, R. Hegde, S. Junuzovic, Z. Liu, C. Zhang, and Z. Zhang, ‘Requirements and Recommendations for an Enhanced Meeting Viewing Experience,’ Proc . ACM Int’ lConf. Multimedia, pp. 539- 548,2008.

[8] K. Otsuka, H. Sawada, J. Yamato, (2007), ‘Automatic Inference of Cross-Modal Nonverbal Interactions in Multiparty Conversations,’ Proc. Int’l Conf. Multimodal Interfaces (ICMI ’07), pp. 255- 262.

[9] M.S. Magnusson, “Discovering Hidden Time Patterns in Behavior: T-Patterns and Their Detection,” Behavior Research Methods, Instruments and Computers, vol. 32, no. 1, pp. 93-110, 2000.

[10] L. Anolli, S. Duncan Jr, M.S. Magnusson, and G. Riva, “The Hidden Structure of Interaction: From Neurons to Culture Patterns,” Emerging Communication: Studies in New Technologies and Practices in Communication, IOS Press, Apr. 2005.