

ANALYZING ORGANIZATION EMPLOYEE MORALE USING INFORMATION GAIN BASED EXPLAINABLE AUTO-ENCODER TECHNIQUE

Harish Bonthu*

Narendrasivam Karthikeyan**

Abstract

Organizational change is a necessity for organizations to survive and prosper in its further growth. Researchers have experimented that employees require continuous change to modify work routines, morale strength and social practices within a firm. E-Mail is considered as one of the important communication medium between managers and employee. However, controlling n number of mail conversation in a single hand is difficult to maintain and understand the relations among the mails. In addition, this process is also difficult for an individual to self-analyze their performance over a period of time. To address these issues, the proposed method designed an Information gain based explainable auto-encoder technique on e-mail social network analysis to compare the communication behavior of managers and employee. This study compared differences in communication patterns by computing social network metrics, such as betweenness, closeness centrality, and content analysis indicators, such as emotionality, complexity of the language used. Results indicate that managers require their degree and closeness centrality with the employees through reduced complexity of their language, as well as their oscillations in betweenness centrality.

Keywords: *Content Analysis Indicators, Communication behavior, E-Mail Social Network, Employee performance, Managers.*

1. Introduction

While the number of communication technologies used in organizations continues to grow [1], e-mail remains the dominant form of communication in the workplace [2]. The way people use the technology is important to study, given the way it influences people's behavior. Technologies such as e-mail therefore influence the way people act, and the technology itself identifies how employees use it have become crucial ways that organizational social systems are created, maintained, and changed [3]. E-mail becomes a critical part of the social context in an organization and "is a key factor in organizational change". E-mail therefore has the potential to empower the powerless to enact change in an organization. Given that premise, the current study examines how e-mail is used to

* Software Architect, Access Technology Solutions

** Technical Architect, Access Technology Solutions

communicate organizational dissent between in-group and out-group organizational members [4,5]. Researchers have been investigating the determinants of employee and managerial turnover for several decades [6]. Factors such as job satisfaction, economic conditions, and personal motivators are among the variables most frequently reported as leading to voluntary turnover [7]. Employees may dissent the organization in one of three ways by engaging in articulated, latent, or displaced dissent.

Although there are many factors that lead to organizational dissent, which may include factors such as the availability of resources, performance evaluations, organizational inefficiency, organizational change, ethical considerations, and worker treatment [8]. There are also many organizational factors that influence dissent, such as employee burnout, employee perceptions of organizational justice, workplace freedom of speech, tenure and so on. Organizational dissent is also influenced by employees' personality traits, including locus of control and aggressive communication traits [9,10]. This study focuses on the ubiquitous use of e-mail in organizations and its effect on employees' willingness to dissent. The research study develop an Information Gain based Explainable Auto Encoder as IGEAE to explain the detected entities, and identify the main features that cause the employees' dissent. This study is embedded into a long history of examining the construct of turnover in terms of relationships. While most of the previous studies [11] have used the intention to leave as dependent variable, we correlate the actual number of managers leaving their job with measures of centrality, responsiveness to e-mail, language complexity and emotionality of the messages. There is a lack of research examining the individual behavior that could lead to managerial turnover. Using a e-mail social network approach to collect and analyze data on communication style, we demonstrate the analytical power of traditional social network metrics such as closeness, betweenness and degree centrality, as well as novel indicators such as response time and number of nudges sent and received by employees.

The remaining of the research paper consists of: Section 2 describes the survey of existing techniques, which are used to identify the employee and managers communication behavior. Section 3 presents the problem statement of the research work. The explanation of the proposed methodology is depicted in Section 4. The validation of proposed method against existing techniques is presented in Section 5. Finally, the conclusion of the research work with its future development are depicted in Section 6.

2. Literature Review

In this section, the survey of recent techniques are discussed, which are used to identify the communication behavior between employee and managers. The advantage of the methods with its limitations are also discussed.

P. A. Gloor, *et al.*, [12] designed a method based on e-mail social network analysis to compare the communication behavior of managers who voluntarily quit their job and managers who decide to stay. The method collected 18 months of e-mail and analyzed the communication behavior of 866 managers. The social network metrics namely, content analysis indicators, betweenness and closeness centrality were computed to compare the difference in communication patterns. The main intention of the work was to educate managers and employees to recognize their own communication patterns before they leave a company and reflect on their position in the communication networks. The experimental results stated that on average managers who quit had lower closeness centrality and less

engaged conversations. The method recognized the intrinsic limitation in using only the subject of a message, instead of the whole e-mail message. In addition, to achieve the high turnover intention, it is important to map the friendships and the organization.

A. J. Nyberg, *et al.*, [13] integrated the fundamental principles of economics and psychology to identify and incorporate employee characteristics, job characteristics, pay system characteristics, and pay system experience into a contingency model of the pay-for-performance (PFP)-future performance relationship. The method used to test the role that these four key contextual factors play in PFP effectiveness using 11,939 employees over a 5-year period. Then, it found that merit and bonus pay, as well as their multiyear trends, were positively associated with future employee performance. In addition, PFP was more effective in jobs where performance can be more easily measured. Finally, the results provided evidence that PFP efficiency must be addressed using both economic and psychological perspectives. The method had some important limitations: Future performance appears more favorably associated with bonus than with merit pay. One potential effect not examined in this study involves supervisor effects. Supervisors may influence the PFP–future performance relationship through a variety of actions.

N. Iqbal, *et al.*, [14] investigated the effect of leadership styles practiced in an organization and their effect on employee performance. The purpose of this study was to understand the effect of different leadership styles autocratic, democratic, and participative style on employee performance. The results stated that leaders only have the authority to take decisions in which employees' felt inferior in doing jobs and decisions in autocratic style. But, employee have to some extent discretionary power to do work in democratic style so their performance was better than in autocratic style. At the end it was concluded that the autocratic leadership was useful in the short term and democratic leadership style was useful in all time horizon. And participation leadership style was most useful in long term and effect on employees was positive.

M. Moussa, *et al.*, [15] composed of several researches about electronic monitoring and privacy concerns, motivations behind employee monitoring and leadership behaviors. The method considered the importance of applying new technologies for monitoring the employee behavior, that should not affect many concerns such as needs, aspirations and privacy. Educating employees about the reasons behind monitoring them, developing a wide range of policies and procedures, and communicating them effectively will be vital to successfully implement a monitoring system. However, the failure of employers to recognize employee rights can cause extensive loss, such as expensive lawsuits, damage to the organization's reputation, and impairment of employee values. Another significant issue raised in this study is that the most effective leadership style to monitor employee performance and behavior through technology remains unidentified.

A. K. Turnage, and A. K. Goodboy, [16] examined whether leader-member exchange status (in-group vs. out-group) of employees explains differences in organizational dissent (i.e., articulated, latent, displaced) through e-mail as opposed to face-to-face. The method considered 166 full-time employees working in a variety of organizations. The results indicated that out-group employees were more likely to express articulated dissent through e-mail, whereas in-group employees were more likely to express articulated dissent in person. The results of this study suggested that the quality of the supervisor subordinate

relationship was important in determining how contradictory opinions were communicated in an upward manner via e-mail. From the study, the major conclusion was that organizations should consider e-mail as a worthy outlet for dissent-one in which useful feedback helped to improve decision making, monitor questionable behavior, and improve relationships in the organization. The limitations in this study included the borderline internal reliability estimates for the latent and displaced subscales for in-person dissent. Another limitation was the small cross-section of employees who work in precision production, craft, and repair, and operators, fabricators, and laborers.

3. Problem Statement

Employee performance includes executing defined duties, meeting deadlines, employee competency, and effectiveness and efficiency in doing work. Various organizations need strong leadership styles that stimulate the employee performance. Some organizations such as tractor factory face the problems: poor innovation, low productivity, inability to meet performance targets. This problem happen due to lack of strategic interventions of specific leadership styles to the particular situations was predicted as the problem at hand. This problem was continuously affecting employee performance. That's why study investigates the best one leadership style that stimulates performance and loyalty of employees.

4. Proposed Methodology

In the recent past years, leadership has engaged as a new effective approach for managing the employees and organization at large. The traditional concept of personnel administration has gradually replaced with the human resource management. This give importance to the strategic integration of new leadership styles into effective management of employees and to improve the employee performance. Different leadership styles are used that fit to employees on the basis of amount of directions, empowerment, and decision making power. An administrative phenomenon reflects the contingency of leadership, and style, situation and performance criteria have been left to suffocate on their own. As a result, employee performance was affected due to lack of proper direction and application of strategic style in managing daily duties.

In past, the previous study investigated performance phenomena and how it was affected by various variables such as: Leadership, and with its different leadership styles such as participative, autocratic, and democratic. It was noted that, in most studies, the concept of employee and managers communication behaviors are not yet considered, especially in large work environments. To address these issues, an Information Gain based Explainable Auto Encoder as IGEAE technique is used to identify the performance of employee via e-mail conversation between managers and employee. Figure 1 shows the block diagram of IGEAE technique.

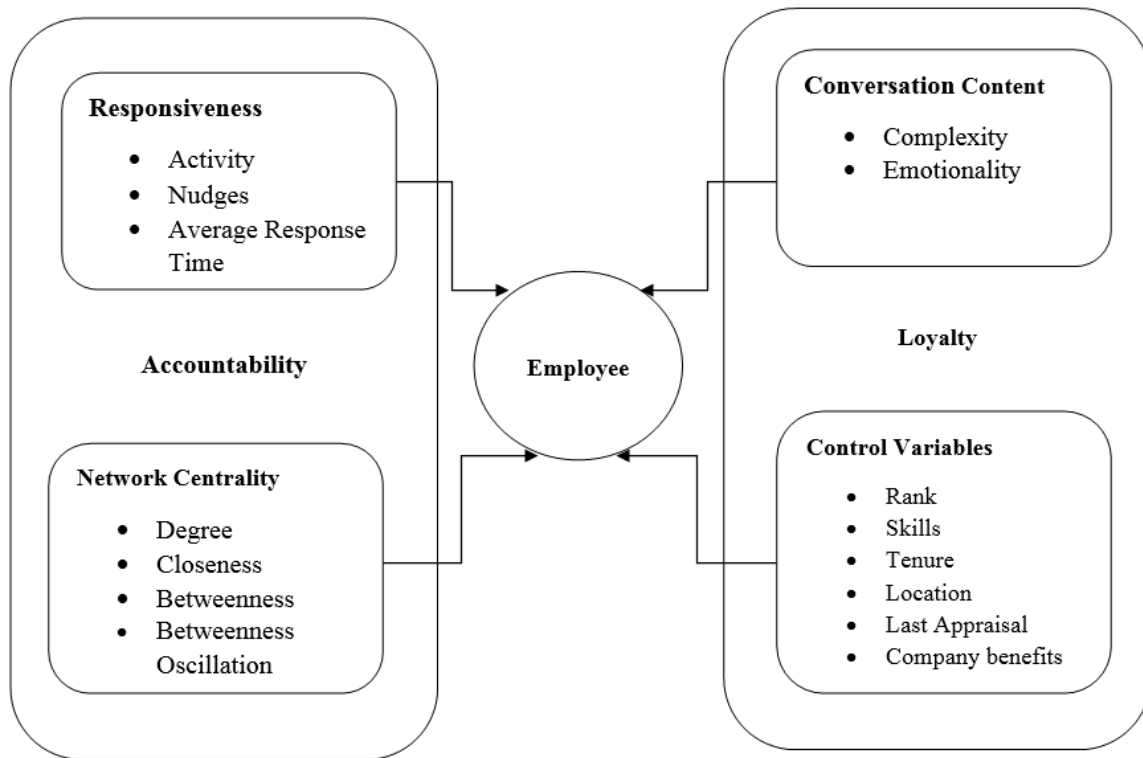


Figure 1: Block Diagram of IGAE Technique

To operationalize centrality, the method adopted three metrics which are well-known and commonly used in the social network analysis literature to identify dominant roles and prominence of actors. An e-mail network can be represented as an oriented graph composed of a set of n nodes (e-mail accounts), which are referred as $G = \{g_1, g_2, g_3 \dots g_n\}$ and a set of m oriented arcs (e-mails) connecting these nodes. Degree centrality considers the number of arcs adjacent to a node, and in our network it represents the number of direct e-mail contacts of an employee. The higher the degree centrality, the higher the number of other people directly reached by that employee.

Betweenness centrality focuses on the capacity of a node to be an intermediary between any two other nodes. This measure is higher when an employee more frequently lies in the indirect communication patterns that interconnect other employees or people external to the company. A network is highly dependent on actors with high betweenness centrality, because of their position as intermediaries and brokers in the information flow. The betweenness centrality of the node g_i is calculated counting the number of shortest paths linking all the generic pairs of nodes and dividing it by the number of paths which contain the node g_i .

The method also monitored betweenness centrality oscillations over time. An oscillation in betweenness centrality indicates that employees shift over time their active involvement in the communication flow, especially their role in transferring information from one person to another. A network with more oscillating leaders is usually more participative and less dominated by few individuals with a stable network position. This study operationalize the measure of betweenness

centrality oscillations counting the number of times a social actor changed his/her score of betweenness centrality (calculated weekly), reaching local maxima or minima, within the time interval of the study.

The other social network metric is closeness centrality, which is based on the average length of the paths linking a node to others and reveals the capacity of a node to be reached, or to reach the others. The more central a node is, the shorter its communication paths. This measure can also be considered as a proxy of the speed. To identify a proxy for the level of engagement within the organization, the method relied on network metrics developed specifically for e-mail networks. In particular, this research method uses the communication activity via e-mail, which indicates the number of e-mail messages sent by a person within a time interval, and nudge, which represents the number of pings (messages) a recipient receives before responding to an email. Also, the differentiate between ego nudges (i.e. number of pings before a recipient responds) and alter nudges (i.e. the number of pings before others respond). Finally, the average response time (ART) is used to measure how much time it takes a person to reply to a particular e-mail. This metric is helpful to identify fast and slow communicators and possibly recognize patterns of behavior looking at periods of slower response.

Using the deep learning algorithms included in the social network and semantic analysis, the method computed other two metrics: complexity and emotionality of the language used. Complexity denotes the deviation of word usage with the assumption that, the more deviate from common, general language, the more complex is our language. Complexity is calculated as the likelihood distribution of words within a message. The method uses the IGEAE for identifying the complexity and emotionality of employee and also used to explain all these parameters. Before explaining the IGEAE method, we provide a short discussion about Auto-Encoder (AE).

4.1. Auto-encoder

An AE is made of three main layers which correspond to (i) the input layer to take in the features, (ii) the latent representation layer of the features, and (iii) the output layer which is the reconstruction of the features. The AE consists of two parts called encoder and decoder respectively. The encoder maps the input into its latent representation while the decoder attempts to reconstruct the features back from the latent representation. The encoder may be deep in the sense that information from the input is passed through several mappings (and hidden layers) similar to the deep architecture in a supervised deep learning model; likewise for the decoder. In this paper, we set the size of the latent representation layer to be 100. In addition, the encoder and the decoder each have three hidden layers with size 512, 512, and 1,024 respectively. The links between the layers show how the values of the next layer can be computed. Commonly, the value of one hidden layer \vec{h}_i can be computed as in Eq. (1),

$$\vec{h}_i = g(W_i \vec{h}_{i-1} + \vec{b}_i) \quad (1)$$

Where, where \vec{h}_{i-1} is a vector of values for the previous layer, W_i is a matrix of weights that signifies the relationship from the previous layer, and \vec{b}_i is a vector of bias terms. Both W_i and \vec{b}_i are parameters to be learned through model training. Here, $g()$ is known as the activation function that transforms the computation in a non-linear way and allows complex

relationship to be learned. Popularly used activation functions include the sigmoid function $g(x) = (1 + e^{-x})^{-1}$ and the rectified linear unit (ReLU) $g(x) = \max(0, x)$, which will be used in this paper. The learning of the parameters are generally achieved by minimizing the reconstruction errors (e.g., mean square errors) via back propagation with random initialization, and can be optimized with a variety of optimizers such as stochastic gradient descent.

4.2. Variational Auto Encoder (VAE)

Unlike AE that deterministically encodes the inputs into their latent representation and subsequently produce a reconstruction, the VAE [17] is a generative model that treats the latent representation layer as random variables conditional on the inputs. Although the encoder and decoder in the VAE follows the same computation model as the AE as in Equation (1), the encoding process is instead used to compute the parameters for the conditional distributions of the latent representation. The parameters can then be used to generate or sample the latent representation for decoding. The conditional distributions are generally assumed to be Gaussian for real-valued nodes. The probabilistic nature of the VAE also means that it cannot simply employ the usual learning algorithm on standard objective function (e.g. mean square error) to train the model. Instead, a class of approximate statistical inference method are used, which is called the Variational Bayes (thus gives rise to the name VAE). The ReLU activation function is used by the encoder and the decoder in all of the intermediate layers, and the linear activation $g(x) = x$ will be used for the output.

4.3. Information Gain based Explainable Auto-Encoder

While most existing works in the literature consider only the evaluation on prediction, the research go beyond and provide an explanation on why a data point is flagged as abnormal. This is significant since it challenges the popular belief that a deep learning model functions as a black box that cannot be interpreted and a VAE model is used. The challenge is to find a method to explain all the entities used in this research work, where existing explainability methods are used for explaining a prediction (output). Here, the study used AE to detect entities through the reconstruction error. The method explain an entity, which is the difference (error) between the input value and the output (reconstructed) value. An entity, if exists, resides in the values of the input and the explaining model needs to explain why this instance is not predicted (reconstructed) well. Thus an explanation must be connected to the error.

This is done by analyzing the gradients ‘contributed’ by each feature of the data point, which is obtainable from the VAE through automatic differentiation in Tensor Flow. Intuitively, given the trained VAE and an anomalous data point, if the function (reconstruction error) changes quite a lot when a particular feature of the anomalous data point is varied by a small amount, then this feature at its current value is significantly abnormal, since it would like to perturb the VAE model (through optimization) to fit itself better. Gradients, or more technically the derivative of the variational lower bound are computed for each feature from each data point. The major applications of the gradient can be immediately derived. Initially, even without having the ground truth labels, the flagged entities can be clustered based on their gradients into groups that share similar behavior, making it easier for analysts to investigate. The entities are detected through the distance computed from the normalized gradient vectors.

The other control variables are used in this research study, which included the manager's internal rank; months since last promotion, that could reduce job satisfaction and increase turnover; tenure within the company and skill (e.g. marketing, supply chain, Information Technology). In next section, the validation of proposed method is given in a brief description.

5. Results and Discussion

In this section, the performance of IGEAE technique is analyzed against existing deep learning techniques by means of several parameters. Initially, the collection of dataset are explained in the below section. Finally, the validated experimental results are given in graphical representation.

5.1. Dataset Description and Research Settings

In this experimental setup, the proposed IGEAE method was implemented on a computer with 8GB RAM, Intel Core i5 with 2.2 GHz using Python 3.7.3. The research setting was a large, global services organization operating in three major cities of South Indian Region, with different software company and more than 150 employees from each company at the time of this study. The method obtained access to e-mail data with the possibility of fetching e-mail messages from the company servers. Each network was analyzed over 10 months, starting from one financial year i.e. April 2015 to March 2016. As a first step of the analysis, the method compared the communication behavior of employee and managers who worked throughout the whole period. The method were not given data on gender and age of the participants; managers retiring naturally were not included in our sample.

5.2. Analysis of Proposed Method

In the advent of employee morale analysis using IGEAE , the system found out the critical factors such as employee turndown, employee performance, morale strength and accountability to be depicted from processing with its corresponding explainability by IGEAE.

5.2.1. Employee Turndown:

Employee Turndown refers to the percentage of probability that an employee can leave the job from the company or try apply for job to any other company. IGEAE relates this factor with the centrality in the network hierarchy of the organization. IGEAE shows this relation in an explainable way for the different ranks of employees like Trainee, Team member, Team Lead and Manager as follows in the figure 2.

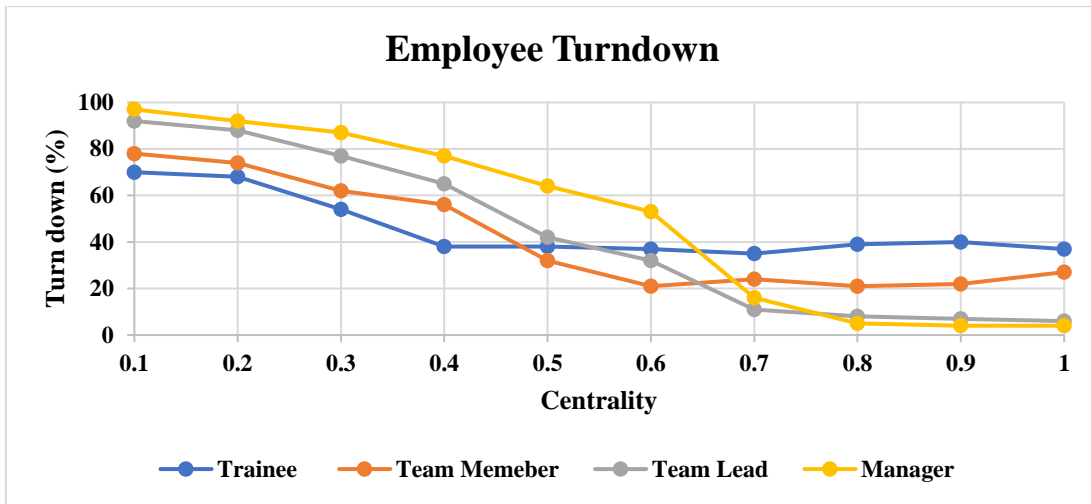


Figure 2. Employee Turndown with respect to network centrality

In figure it is observed that the trainee and team member has a lesser percentage of turndown when network centrality is less and getting decrease further when trainee or team member has the significant responsibility around 50%. This turndown remains as oscillatory in the same range even if centrality increases due to their common increasing mentality of expectation for a promotion or an appraisal for their increased centrality responsibility. Whereas in the case of team lead or manager, if the centrality is down with their respect to their position they feel humiliated and their turndown percentage increases. When centrality and responsibility increases team lead and manager gets satisfied and their loyalty increases with the reduction in turndown to very minimal state.

5.2.2. Employee Performance

Employee performance can be measured by IGEAE using the amount of deliverable created by the employee or the amount of tasks completed within a working day corresponding to their position and skills. Hence, IGEAE describes the employee performance in relation to the number of activities done per working day with 8 hours per day as like in the depicted figure 3 below.

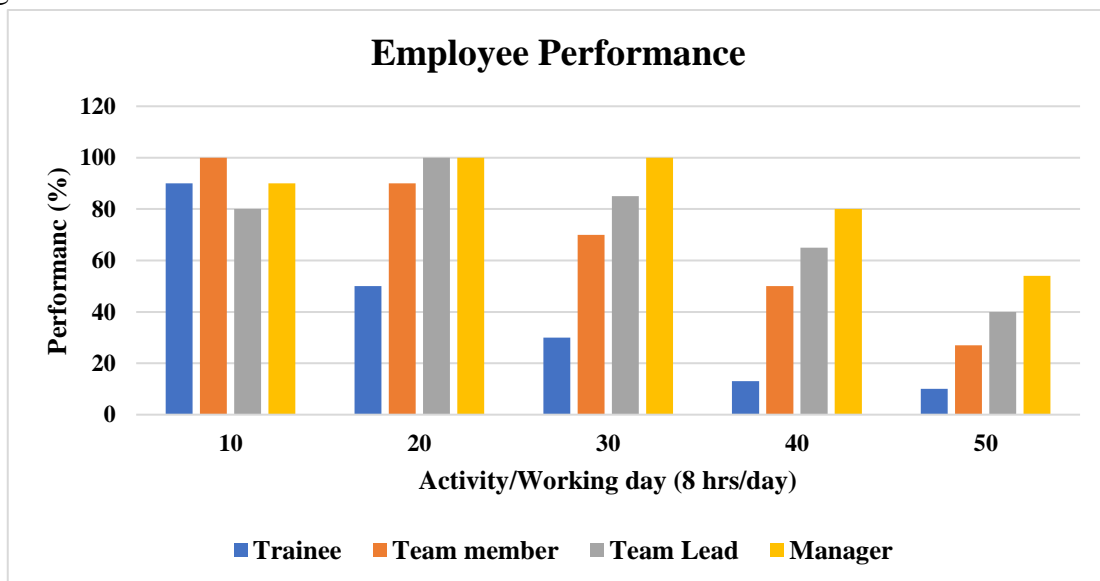


Figure 3. Employee performance based on activities assigned per day

In figure the employee performance is depicted by IGAE with relation to the number of activities done assigned to the employee. In this diagram it shows that the performance of trainee and team member is better during less works and getting reduced when their activity ranges more than 20 per day. Meanwhile the team lead or manager expecting to be involved in more activities so that they can delegate the work to others and increase the profit so their performance getting reduce in lesser activities and maximized at average. In all the cases the performance is getting much reduced when the activities became an over burden to the employees.

5.2.3. Employee Morale

IGAE relates the employee morale with the factors of conversation complexity and emotionality for the activities done per working day. Based on how the task is being comprehended by the employee and the emotion expressed by the employee towards the conversation. It is being reflected in the number of tasks that employee can handle at a time. This can be depicted from the below figure 4.

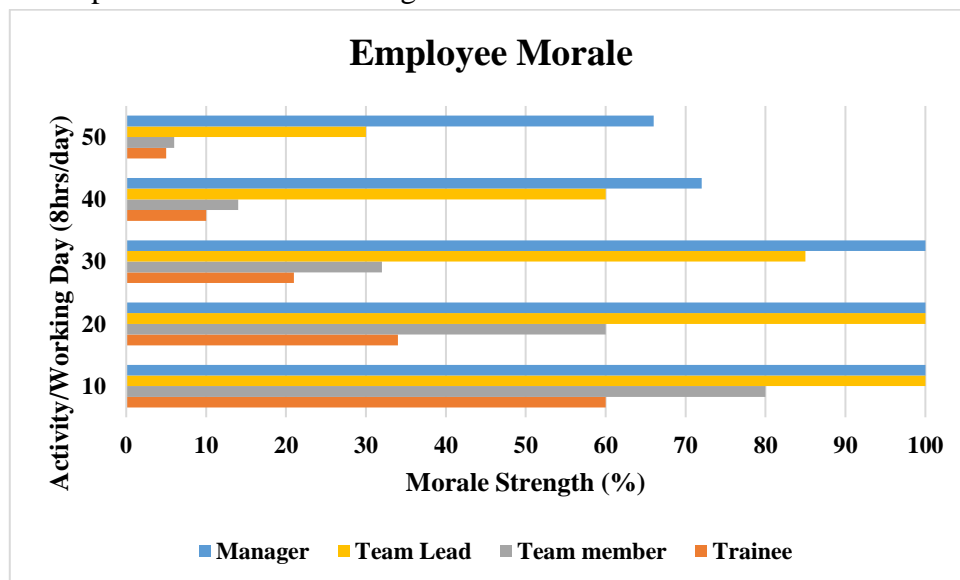


Figure 4: Employee Morale with respect to tasks

Employee Morale is found to be lower for a trainee and team member in all cases because of their insecurity in job and inexperienced nature towards approaching the tasks. This reflects their morale to be average in cases of less tasks and very lower when tasks increasing to higher extent. In case of a team lead or manager their morale is very high because of their high experience and skills in situations of lower and average tasks. When it goes for a higher tasks managers remain calm in morale when compared to team leads because of their problem solving skills. When tasks getting over burden in general all the employees losing their morale because of pressure.

5.2.4. Employee Accountability

The proposed IGAE method, makes the relation for company's accountability towards employee in respect to the betweenness oscillation employee possess in the company's hierarchical network. Amount of interest the employee possess for the work and benefit of organization is directly related to the responsiveness of employee and the amount of imposed factors done by the company through its control variables like position, rank and financial

or personal assistance given by company policies. This can be shown through the following figure 5.

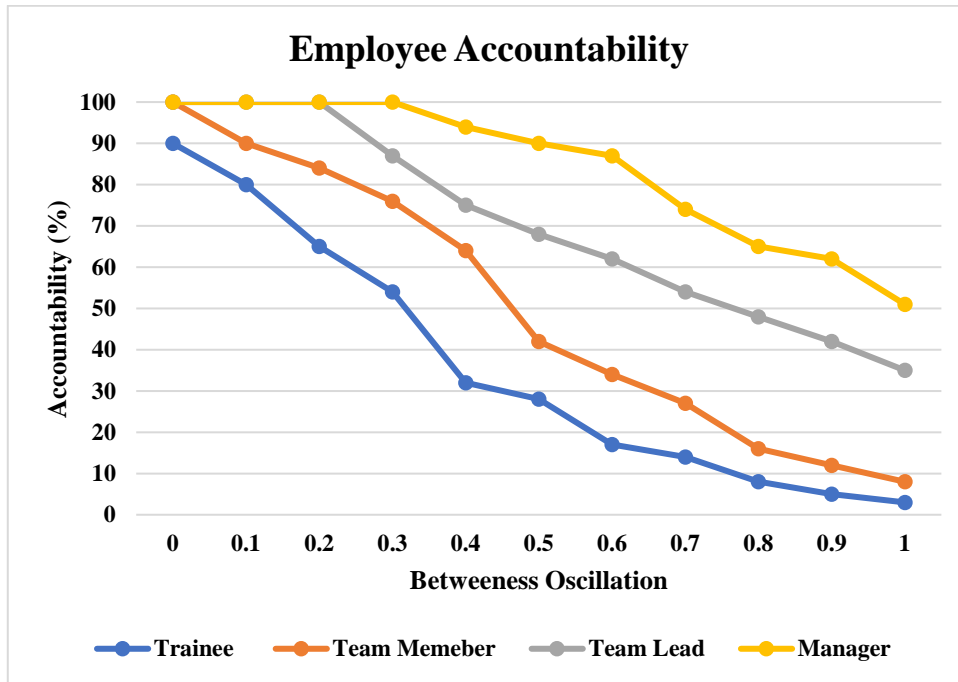


Figure 5: Employee accountability through betweenness oscillation

In IGAE, the employee's betweenness oscillations corresponds to the employee's lack of interest in the work because of this reason company's losing accountability on the employee. In otherwise the employee loses trustworthiness due to lack of participation in the given role. Generally, this accountability drop happens as a sudden fall when it crosses 0.4 to 0.5 betweenness oscillation for a trainee or a team member because of the excuses permitted in earlier stages by higher officials and sudden loss of hope on the employees on later stage. But in case of a team lead or a manager, they possess more accountability till 0.7 to 0.8 betweenness oscillations due to the case they can't involve in all low level conversations and involvement requires only in critical decision making places. Team lead and manager will lose their accountability to a maximum stage up to 50% only even in worst cases.

6. Conclusion

Using e-mail communication analysis, along with explanation deep learning method to assess employees' satisfaction, human resource managers can offer the most appropriate organizational initiative such as mentoring programs, cross staffing, or communities of practice that leverages the need for interdisciplinary efforts. The application of the social network method described in this paper has the potential to help managers to better understand the nature of managerial turnover at their particular organization. In addition, the proposed method offers an opportunity to begin a self-reflection and plan proactive strategies that are tailored to the overall patterns observed in the past. Finally, the analysis of this research want to recognize that monitoring organizational emails carries some ethical issues that should also be considered before conducting a similar research. The major intent is to educate managers and employees to recognize their own communication patterns before they leave a company and reflect on their position in the communication networks. The proposed method's concern is with protecting the individual respondents, and confidentiality and anonymity should be clearly embedded into any social network interventions.

Reference

- [1] D'Urso, S. C., & Pierce, K. M. (2009). "Connected to the organization: A survey of communication technologies in the modern organizational landscape". *Communication Research Reports*, 26, 75-81.
- [2] D. D. Dawley, and W. P. Anthony, (2003). "User perceptions of e-mail at work". *Journal of Business and Technical Communication*, 17, 170-200.
- [3] Katz, J. E., & Rice, R. E. (2002). "Social consequences of Internet use: Access, involvement, and interaction". *Cambridge, MA: MIT Press*.
- [4] Stephens, Keri K., Renee L. Cowan, and Marian L. Houser. "Organizational norm congruency and interpersonal familiarity in e-mail: Examining messages from two different status perspectives." *Journal of Computer-Mediated Communication* 16.2 (2011): 228-249.
- [5] Van den Hooff, Bart, Jasper Groot, and Sander de Jonge. "Situational influences on the use of communication technologies: A meta-analysis and exploratory study." *The Journal of Business Communication* (1973) 42.1 (2005): 4-27.
- [6] Holtom, B. C., Mitchell, T. R., Lee, T. W., & Eberly, M. B. (2008). "5 turnover and retention research: a glance at the past, a closer review of the present, and a venture into the future". *The Academy of Management Annals*, 2(1), 231-274.
- [7] Egan, Toby Marshall, Baiyin Yang, and Kenneth R. Bartlett. "The effects of organizational learning culture and job satisfaction on motivation to transfer learning and turnover intention." *Human resource development quarterly* 15.3 (2004): 279-301.
- [8] Kassing, Jeffrey W., and Todd A. Armstrong. "Someone's going to hear about this: Examining the association between dissent-triggering events and employees' dissent expression." *Management Communication Quarterly* 16.1 (2002): 39-65.
- [9] Carter, M. Z., Armenakis, A. A., Feild, H. S., & Mossholder, K. W. (2013). "Transformational leadership, relationship quality, and employee performance during continuous incremental organizational change". *Journal of Organizational Behavior*, 34(7), 942-958.
- [10] Budworth, Marie-Hélène, Gary P. Latham, and Laxmikant Manroop. "Looking forward to performance improvement: A field test of the feedforward interview for performance management." *Human Resource Management* 54.1 (2015): 45-54.
- [11] Soltis, S. M., Agneessens, F., Sasovova, Z., & Labianca, G. (2013). A social network perspective on turnover intentions: The role of distributive justice and social support. *Human Resource Management*, 52(4), 561-584.
- [12] P. A. Gloor, A. F. Colladon, F. Grippa, and G. Giacomelli, "Forecasting managerial turnover through e-mail based social network analysis". *Computers in Human Behavior*, 71, 343-352, 2017.
- [13] A. J. Nyberg, J. R. Pieper, and C. O. Trevor. "Pay-for-performance's effect on future employee performance: Integrating psychological and economic principles toward a contingency perspective," *Journal of Management* 42.7, 1753-1783, 2016.
- [14] N. Iqbal, S. Anwar, and N. Haider. "Effect of leadership style on employee performance." *Arabian Journal of Business and Management Review* 5.5, 1-6, 2015.
- [15] M. Moussa, "Monitoring employee behavior through the use of technology and issues of employee privacy in America." *Sage Open* 5.2, pp. 2158244015580168, 2015.

- [16] A. K. Turnage, and A. K. Goodboy. "E-mail and face-to-face organizational dissent as a function of leader-member exchange status." *International Journal of Business Communication* 53.3, 271-285, 2016.
- [17] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," arXiv preprint arXiv:1312.6114, 2013.