RESEARCH ARTICLE OPEN ACCESS

Reducing Churn Rate Using Business Intelligence Approach in Data Mining

Trasha Gupta[1]
Research Student[1] Department of Computer Science Engineering,
Deen dayal Upadhyaya College, University Of Delhi,
Karampura -Delhi, India

ABSTRACT
Understanding and analyzing the factors influencing future participation of its employees is relevant to an organization. This information is useful for planning and strategic decision-making. In this work, we analyze contributor characteristics to reduce the attrition by investigating their relationship in mining Issue Tracking System. We conduct experiments on four year data extracted from Google Chromium Issue Tracking System. We investigate the factors for future participation in projects.
Keywords:- Churn Rate, Component-Developer model, Developer-Developer model

I. INTRODUCTION

The success of any organization to a very large extend depends on the employees it retains, which in turn depends on the organizational structure and policies. Among the various influencing factors, attrition rate plays a key role, which may be defined as the reduction in workforce due to retirement, resignation or death. The two factors 'retirement' and 'death' are inevitable. So, in the following work we focus on the impact of attrition rate on the IT industry with emphasis on Technical Engineers in particular. Here an analysis of the reasons and possible solutions will be discussed. The solution discussed here will be of concern to the end-user i.e. the HR manager.

A. Basic Terminology

Basic terminologies used in this paper are defined below

Attrition rate: It has been defined as the reduction in workforce due to retirement, resignation or death. The two factors 'retirement' and death are inevitable.

Assigned_to vs Technical Expertize model: This model captures the way in which people work on bugs with different technical expertise. Some people will be found working most on one topic, while others are seen making their hands dirty in variety.

Assigned_to vs CC model: This model captures the people to whom any person turns down most for help. These are the people added in the cc-list of that bug.

CC vs Technical Expertize model: This model provides a validation/inferential data for the observations made under the following head. The technical expertise of the individual is compared with that of their most involvements.

II. MOTIVATION

The main motivation for carrying out the research is industrial success in Japan. This research is suggested to reduce the attrition rate in industries for India. The highest attrition rate of 23% in the first quarter of 2010-11 witnessed by IT, ITES sectors in India. [3]

A. From a business perspective

The following lists some factors resulting from attrition from business perspective.

- When a business loses employees, it loses its employees; it loses skill, experience and corporate money. The magnitude and nature of these loses is a critical management issue, affecting productivity, profitability, and product and service quality.
- The cost of replacing workers can be high, the problems associated with finding and training new employees can be considerable, and the specific workplace-acquired skills and knowledge people walk away with, can take years to replace
- Following additional costs incurred due to turnover
  a) Recruitment cost
  b) Training cost
  c) Lost Productivity cost
  d) New Hire Cost
  e) Lost sales cost
- As has been stated by Deloitte Global Service Line Leader (Global Employer Services) Bruce M Steward, attrition cost is about three or four times of the individual's salary.
- The arising risk for the organization calls for talent retention initiative on the part of businesses.
- Attrition rates may go up by 25% in 2011.

B. From an employee's perspective

From employee's perspective, it can lead to high turnover which can negatively affect employment relationship, morale and workplace safety.

C. Impact in India

The following reasons have been identified as the cause of switching.

Pay Packages	21%
Career level Growth	16%
Dissatisfaction with supervisors	15%
Work pressure	14%

III. RELATED WORK

The concerns for business success have always called up for ways to mitigate the factors which hampers the growth of any organization. In the past, a lot of study has been done to identify the causative factors. One of those is attrition rate or turnover. These factors are very closely related to the most important asset for an organization i.e. human being. A study by Pratibha et.al talks about the following factors.

- Higher Pay Package in another company,
- Good working Conditions, in new company,
- Opportunities for growth in new company,
- Change of Place problem,
- A better Boss in new company,
- Brand Image of the new company,
- Mismatch of Job Profile in existing company,
- Lack of Empowerment and Employee Engagement,
- Stress and Work-Life Balances,
- Odd Working Hours,
- Lack of Trust.

In this current work, we talk about the job dissatisfaction which directly impacts the above highlighted factors.

IV. PROBLEM STATEMENT

The success of an organization depends on both market and non-market factors. With the growth of social media now people can share their thoughts openly and freely with the masses where the masses (connections) can further participate on the same via their likes/comments (as in Facebook). Also the credibility of the participants varies with the organizational personnel being the most influential. These non-market parameters account for the reputation of an organization. One such non-market parameter is the churn rate (also termed as attrition rate). This parameter symbolizes the employee satisfaction.

V. METHODOLOGY

A. System Environment

The given setup operates in the IT industry (specifically in service based). Roughly the services can be roughly classified under three heads. The top layer constitutes the managerial section that control and coordinate the tasks. The bottom level includes the clerical section who is engaged in serving the middle level workers which is the building block of any organization. This middle layer shares the largest area in the pyramid which is precisely our focus. It constitutes the Technical Engineers, which is the subject of our study.

B. Assumption

We assume that the company holds the wide variety of projects that may be of interest to the population under consideration.

C. Method Used

We conduct a series of experiments on Google Chromium project dataset. Google Chromium is a popular and widely used open-source browser. The Issue Tracker for Google Chromium project is publicly available and the data can be downloaded using Google Issue Tracker API. Following is the list of attributes identified in a typical bug report extracted from a Chromium data set.

- **IssueIDA:** unique identifier for each bug report
- **State:** Closed or Open.
- **Status:** Fixed, Duplicate, and Verified
- **Reported Time-Stamp**: The starting and closing time of a bug
- **Reporter:** The person who reported the bug
- **owner:** The person who is assigned the bug
- **Title:** The label of the bug
- **Description:** Long description of the bug
- **Area:** The intuitive area of bug. It can have values: UI, internals, webkit, build, compat, chrome frame etc.
- **Type:** The nature of bug. It can be Bug, Feature, Regression, Usability, Localization, Polish, Other, Security, Task, Clean-up, Meta, Test, Documentation, Defect, Feature Tracker, Yak
- **Priority:** 0,1,2,3
- **Length of Description:** Total length of the description
- **Operating System:** The operating system on which bug is observed.

Relevant relationships and models (as discussed in section I) were built using MySQL. These models were visualized and analysed using ORA. Below presented are the models generated.

Initial View: The isolated vertices represent the people who made no contribution to the bugs with status RESOLVED and resolution FIXED. Our focus in this work is the strongly connected component which is present at the core of the Figure shown below.

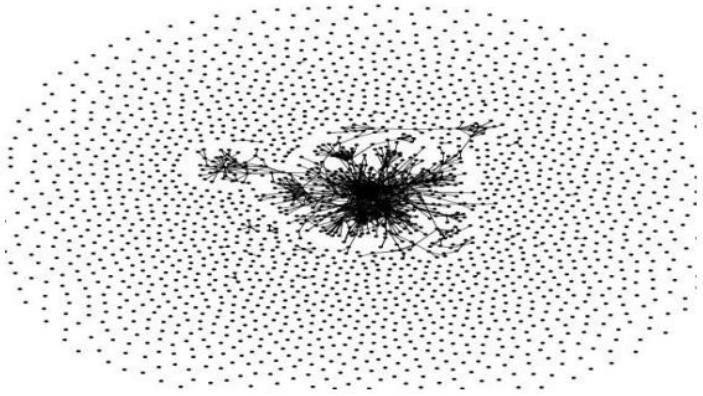

Figure 1: Initial View of the population

Component-Developer Model (Assigned_to Vs Technical Expertise model): In figure 2, the green nodes represent the component. The red nodes represent the developers. As can be seen from the figure, the working Patterns of technically expertise people is to a great extend confined to some particular technical expertise. These are the people who have worked a lot on any particular domain for instance API. Figure 3 represents network with link weight greater than 50.

Developer-Developer Model (Assigned_to vs CC model): This model tracks the collaboration pattern of experts of domain i.e. the people to whom the bug was assigned initially to those who were kept in the CC-list .The figure 4 shows the highly collaborative Patterns amongst the two group's i.e whenever some person in assigned_to Works on a particular domain, it is seemly found that some particular person was there in the cc_list or say the persons in cc list was very much interested in the ongoing bug resolution. Our aim was to find the patterns in which people collaborate extensively with the actual resolver of the component because of his/her impressive expertise.

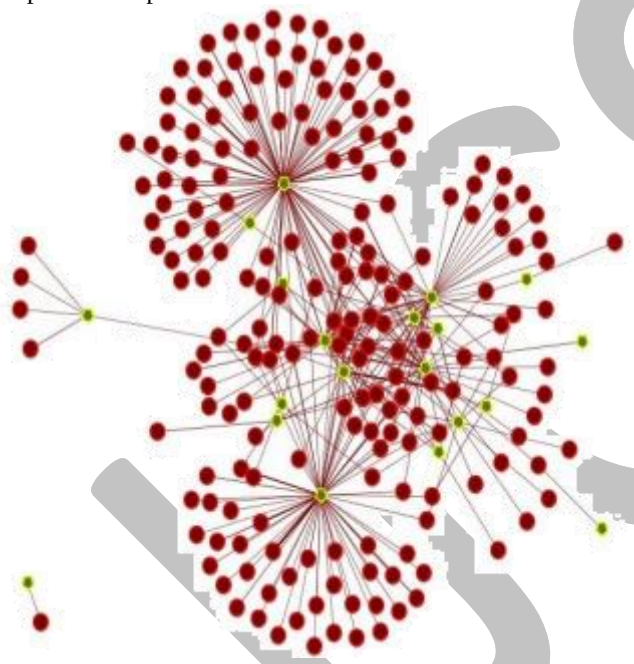

Figure 2: Component-Developer Model I

Component Developer Model II (CC_list vs Technical Interest): This model tracks the collaboration of pattern of domain of interest i.e. the people to whom the bug was not assigned but they were interested in the particular technology. The figure 5 shows the highly collaborative Patterns amongst the people in cc list and the particular component. Our aim was to find the patterns in which people are interested in certain component and assign the component (the domain of interest) to reduce the employee attrition rate.

VI. ANALYSIS

- Agent 51 has been assigned some task but he was more interested in component "performance" so he was later assigned that component. Agent 18 has been assigned performance and but is collaborating on accessibility and later agent 51 joined. Comparing 18's output on performance component with 51 he has been removed from there.

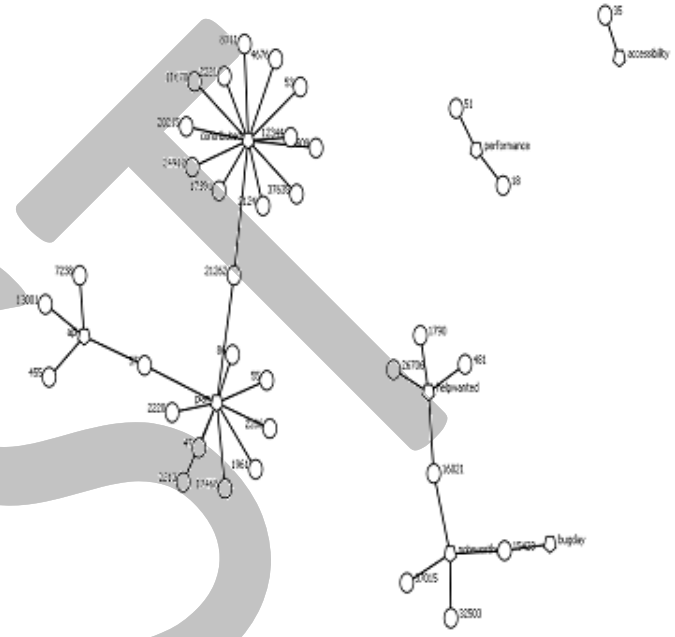

Figure 3: Component-Developer Model with link weight greater than 50

- People like 32503 and 15423 were very much interested on component "noteworthy" and they were assigned the component. Similarly, for agent 8011 was interested in component contributed.
- Some people have been working according to their interest. For instance agent 2124 has been assigned component contributed
- 57 was initially assigned component "plan" but have been continuously collaborating on component API so assigned the API component and has performed well.

A. Limitations

The system environment that we build is based on assumptions. In real world, various other factors must be taken into considerations. It just provides a prototype for building such a network. Currently, its operability is confined to Open Source System (OSS), where people enjoy the freedom to work as per their choice.

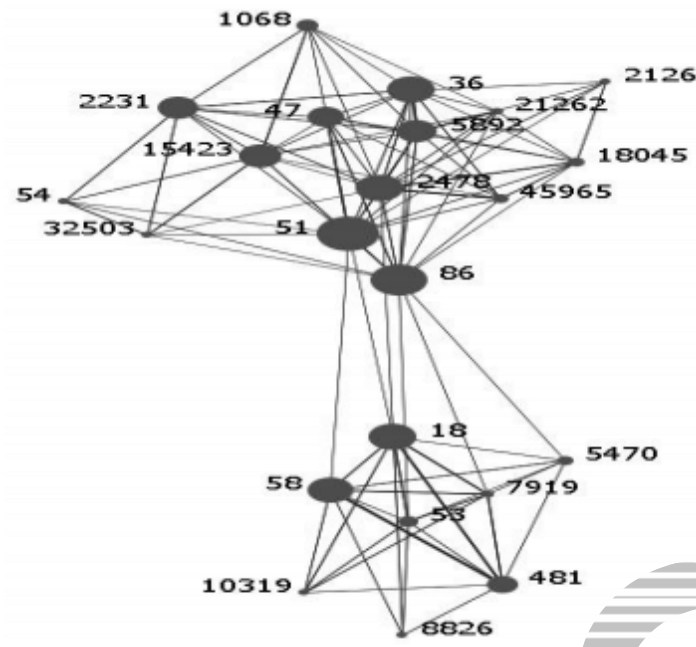

Figure 4: Developer-Developer Model

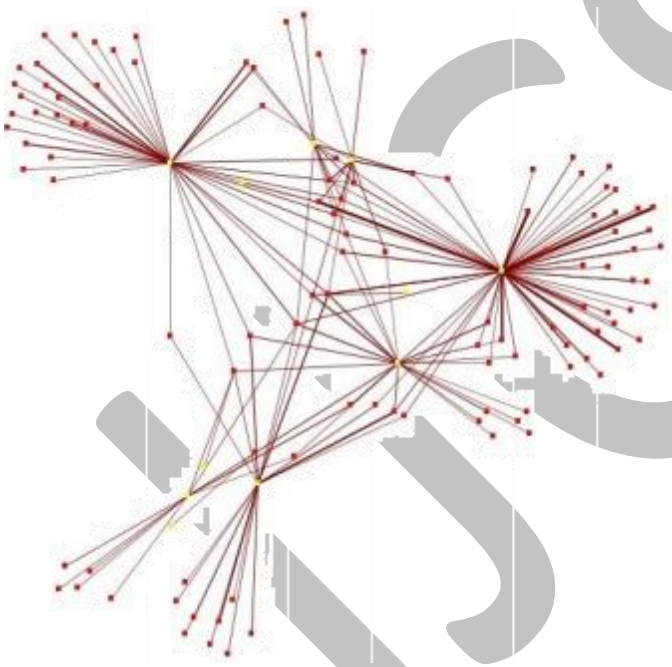

Figure 5: Component Developer Model II (CC_list vs Technical Interest)

VII. CONCLUSION, SCOPE AND FUTURE WORK

A more empirical approach to identify the interest based on work experience and inclination. It is expected to give better results than ad hoc work allocation to the employees. This approach can be very well extended to any sectors of commerce with appropriate attributes and environmental extension. The crux of informal relations (social dynamics) captured to influence the formal relations. Extension of appropriate strategies to the Closed Source System (CSS), the section which dominates the business section.

REFERENCES

[1] Mei-chi Hu, Connie Zheng, David Lamond, 2007, Recruitment and retention of ICT skills among MNCs in Taiwan, Chinese management studies, vol. 1, no. 2, pp. 78-92.
[2] Harvard Business Review, Behind Japans Success, http://hbr.org/1981/01/behind-japans-success/ar/1
[3] http://articles.economictimes.indiatimes.com/2011-07-19/news/29790871_1_attrition-rate-cent-salary-hike-sectors
[4] http://pioneerjournal.in/conferences/mapping-for-excellence/8th-national-conference/3460-job-satisfaction-an-effectivetool-for-reducing-attrition.html
[5] http://www.rediff.com/business/report/high-attrition-rate-hurting-india-deloitte/20110503.htm
[6] [http://www.bpoindia.org/research/attrition-rate-big-challenge.shtml
[7] http://www.youtube.com/watch?v=N-ebIGpZIWI
[8] http://www.youtube.com/watch?v=9SCxJTmS1EoS. M. Metev and V.