Employee Turnover Prediction Based on State-transition and Semi-Markov- A Case Study of Chinese State-owned Enterprise

Ming FANG, Jia-Hao SU, Tao WANG, and Ren-Jie HE

College of Information System and Management, National University of Defense Technology, Changsha, China
*mingxtz@gmail.com, *suijiaha15@nudt.edu.cn, *wangtao@nudt.edu.cn, *839434993@qq.com
*corresponding author: wangtao@nudt.edu.cn

Abstract: As a main direction of Human Resource Management, employee turnover can provide decision support for managers. In this paper, we aim at predicting the turnover amount of employee on condition of different variable values. The properties of employee and job position are formulated as two variables, where the value of variable varies according to the the state of properties. Additionally, state-transition model is applied to describing employee’s job-state as well as the turnover type. Subsequently, we proposed a semi-Markov model to calculate the conditional turnover amount of employee. Then, we provide a dataset of employee records to illustrate how these models work in reality. Finally, it is proven that the proposed method in this paper is with great significance for managers to develop recruitment plans, promote rules, and retire regulations.

1 Introduction

Employee turnover is the leaving act of employee in a workforce. In the study of improving employee satisfaction and decreasing the staff cost for workforce such IT [1,27], tourist [2] and hospitality [3], employee turnover is of particular importance for managers to predict organization structure and develop employee transfer plan.

Employee turnover has long been the focus of Human Resource Management (HRM). Turnover management can provide decision support of managers. Actually, before the turnover behavior happens, an employee first forms an intention to leave the organization. Employee turnover intention is a conscious and deliberate willingness to leave one’s formal organization of employment [4,5]. Various potential factors may have influence on turnover intention, then contribute to turnover behavior.

Turnover intention always comes from individual factors [6,7], organizational factor [8,9] and environmental factors [10,11]. Individual perception is an important factor to the turnover phenomenon, which is a subjective measure of oneself, e.g., how easy it would be to find a new job [12]. The inducement-contribution balance developed from organizational equilibrium theory is touted as the one of the most influential factors of turnover. Employee would like to stay in current position only if the inducement of this organization is greater than or equal to the contribution asked by managers. [4,13] In addition, employee turnover is often influenced by economics, the rate declines during economic recessions [14], while it increases during economic rebound [15].

Afterwards, researchers start to find the connection of turnover intention with turnover behavior [16-18], and the result suggests that the longer the elapsed time between the measurements of intention and behavior, the lower is the correlation [19].

Based on turnover intention, extensive studies have been made on modeling and predicting the turnover. For instance, [20] studies twenty variables to build a turnover prediction model, and only six variables are found effective by stepwise logit analysis; [21] provides a dataset of life insurance company to find the effects of prior absenteeism and performance rating on voluntary turnover; [28] propose a forecast model based on neural network to predict the impact of employee turnover on the performance; recently, machine learning has been applied to turnover prediction by [22].

The turnover can cause serious loss and difficulty because of the leaving of employees who have relatively high human capital [23]. Thus HRM strategies are developed to facilitate the retention of employees [24-26]. Nevertheless, little attention has been paid to Chinese enterprise as well as the internal transfer within organization. In this paper, we propose employee turnover as the combination of inter-organization turnover and internal transfer, with inter-organization turnover denoting recruit, retire, resign and disposal, while internal transfer denoting training on current job position, exchange to fellow position and promote to higher position. In addition, we propose state-transition model to describe the properties related to employee turnover, and employee are classified to deferent groups according to their state properties. Subsequently, conditional semi-Markov (CSMK) model is proposed to predict the turnover on condition of different constraints. Based on CSMK, transition amount can be calculated for each turnover type, then the employed amount can be predicted for each employee group. Lastly, a case study of Chinese state-owned enterprise is provided to illustrate how state-transition and CSMK work in reality. These models are proven to be available to predict the structure of an organization, which is with great significance for managers to develop recruitment plans, promote rules, and retire regulations. This turnover prediction method can be
2 Modeling Based on State-transition Mode

Turnover of employee always contributes to the change of the employee property and the organization structure. Make an assumption that there is an employ-position system that contains both the property of employee and job position. Then, we can define job-state to describe the state of this system. The job-state will transfer to another state if system property changes, because of employee turnover, and this process can be described by state-transition to classify the turnover type. Therefore, we claim that the state-transition mode can be applied to predict the employee turnover.

2.1 Formulating the Properties about Employee Turnover

Job position are classified into i grades according to their authority and function; the bigger the grade is, the higher the position level is. Let a and b be the property of job position and employee respectively; both a and b are related to employee turnover, where \( a=(a_1,a_2,\ldots,a_n) \) and \( b=(b_1,b_2,\ldots,b_m) \). For a job position, \( a_i \) denotes the \( i \)th property of \( a \), and \( a \in A \). Similarly, \( b_j \) denotes the \( j \)th property of \( b \), and \( b \in B \).

Subsequently, \( a_i \) and \( b_j \) are developed to describe states that contain all property values of a position or an employee. Position states are distinguished by their subscripts \( j \) which denote the sequence number of position states, where \( a_i=(a_{i1},a_{i2},\ldots,a_{in}) \) and \( a \in A \). Similarly, employee states are distinguished by \( k \), \( b_j=(b_{j1},b_{j2},\ldots,b_{jm}) \) and \( b \in B \). 

Then, \( \rho \) is proposed to measure the property of manning quotas. As is shown in (3), \( \rho \) describes 4 position states about employee turnover.

\[
\rho = \begin{cases} 
0 & \text{on guard} \\
1 & \text{overstrength} \\
2 & \text{vacancy} \\
3 & \text{opening & overstrength}
\end{cases}
\]

Based on the definition of these properties, we can define the employee job-state as a triple \((ijk)\), or a tetrad \((ijk\rho)\), to describe state-transition mode. The system state will be changed when there is any transition happens to position state or employee state. It should be noted that the total number of \( I, J \) and \( K \) are \( I, J \) and \( K \), respectively.

2.2 Classifying State-transition of Employee’s Turnover

In this paper, state-transition is a formal expression of employee’s turnover, describing the transition among job-states. Without considering the property of manning quotas \( \rho \), state-transition of employee can be described by three-dimensional state \((ijk)\), then all these states can make up a set, \( \Omega \), whose element corresponding to employee group. This group is denoted by \( U_{ijk} \), to group employees with different state-transition. Similarly, the employee’s turnover of retire, disposal, resign, recruit and training are denoted by \( \Omega^\rho \), \( \Omega_r \), \( \Omega_s \), \( \Omega_k \), and \( \Omega_r \), respectively.

\[
\Omega = \Omega^\rho \cup \Omega_r \cup \Omega_s \cup \Omega_k \cup \Omega_r \cup \Omega_s
\]

Then, take \((ijk) \rightarrow (ijk')\) as the employee’s transition from current state to objective state. This transition process can be formulated as (5). Note that \( i'=i \) or \( i'=i' \) when employee is promoted gradually or leapingly; \( j'=j \), when employee exchanges to another position that with the same property as current position; \( k'=k+1 \) after training on current position, without any transition of \( i \) and \( j \); \((ijk) \rightarrow (ijk')\) after recruit of new employee, initializing the job-state. Specially, the transition is denoted as \((ie)\rightarrow (ijk)\) when an employee is recruited as a freshman, where \( e \in \{0,1,2,3\} \) denotes the employee’s education degree is bachelor \((e=1)\), master \((e=2)\) or doctorate \((e=3)\), or else \((e=0)\).

\[
(ijk) \rightarrow (ijk')
\]

While manning quotas is taken into consideration, the transition can be denoted as \((ijk\rho) \rightarrow (ijk\rho')\). These four-dimensional states \((ijk\rho)\) makes up the set, \( \Omega_{ijk\rho} \). The transition of \( i, j \) and \( k \) follows the same rule as the three-dimensional state \((ijk)\). What should be noted is that manning quotas transfers between on guard \((\rho=0)\) and vacancy \((\rho=2)\), or between overstrength \((\rho=1)\) and open & overstrength \((\rho=3)\). \( \rho \) equals 2 or 3 when employee leave his/her current position where \( \rho \) equals 0 or 1, meaning that these transitions provide an open job-state \((\rho=2,3)\) for other employee to replenish while the objective state \((\rho=0,1)\) becomes busy for its occupied by this employee, such as retire, disposal and resign. As is
demonstrated in the following table, \(\rho'\) equals 0 or 1 when employee is recruited, and \(\rho'\) maintains 0 or 1 after employee’s attending training.

### Table 1. Table Type Styles

<table>
<thead>
<tr>
<th>((ijk)\rho)</th>
<th>(\rho)</th>
<th>(\rho')</th>
<th>Transition</th>
<th>Open</th>
<th>Busy</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i'j'k'\rho)</td>
<td>0/1</td>
<td>2/3</td>
<td>promote</td>
<td>(ij'\rho')</td>
<td>(i'j'\rho)</td>
</tr>
<tr>
<td>(ij'k'\rho)</td>
<td>0/1</td>
<td>2/3</td>
<td>exchange</td>
<td>(ij'\rho')</td>
<td>(ij'\rho)</td>
</tr>
<tr>
<td>(ijk'\rho)</td>
<td>0/1</td>
<td>2/3</td>
<td>retire/deposal/resign</td>
<td>(ij'\rho')</td>
<td>--</td>
</tr>
<tr>
<td>(ijk\rho)</td>
<td>0/1</td>
<td>0/1</td>
<td>training</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(ijk'\rho)</td>
<td>2/3</td>
<td>0/1</td>
<td>recruit</td>
<td>(ij'\rho')</td>
<td></td>
</tr>
</tbody>
</table>

\(\rho \geq 0\) is the CSMV process of employee turnover, with variables of \(r\) and \(c\).

\[
\sum_{\omega \in \Omega} P_{\omega \rightarrow \omega} |r,c| = 1 \quad (7)
\[
F_{\omega \rightarrow \omega}(t| r,c) = \Pr(t_{\omega} - t_{\omega} \leq t | X_{\omega}, \omega, X_{\omega}, \omega, \tau, c). \quad (8)
\]

### 3.2 Conditional Calculation of State-transition

Based on the definition of CSMV process, we propose (9), function \(T_{\omega \rightarrow \omega}|r,c\), to demonstrate how transfer probability changes over time. It’s apparent that employee’s turnover depends on transfer probability among states as well as the distribution of interarrival time. Equation (10) and (11) reveal the connection of \(T_{\omega \rightarrow \omega} |r,c\) and \(F_{\omega \rightarrow \omega} |r,c\) when \(T_{\omega \rightarrow \omega} |r,c| = 0\) on condition of \(P_{\omega \rightarrow \omega} |r,c\). Then, \(T\) can be derived as (12), the multiplication of \(P\) and \(F\).

The transfer of employee’s job-state over time can be abstracted as a stochastic process, \(\{X(t), t \geq 0\}\). In this process, \(r\) and \(c\) are provided as variables to classify the job-state of employee, \(r=(r', l)\). The adjoint variable, \(\tau\), is provided to describe employee state, where \(r'\) is a natural factor denoting properties such as gender and native place, while \(l=(l_{r'}, l_{r}, l_{r})\) is an adjoint factor, whose elements, \(l_{r'}, l_{r}\) and \(l_{r}\), denote the age, working age and grade age at the beginning of current job-state, respectively. Another variable is \(c\), the environment to describe properties of job-state group \((U_{ijk})\), such as age, education degree and manning quotas.

Let \(\omega'\) be the objective state of \(\omega\) for employee's turnover, \(\omega \in \Omega\). Thus the probability of state-transition from \(\omega\) to \(\omega'\) can be denoted as \(P_{\omega \rightarrow \omega'} |r,c\), and (7) is the constraint on condition of \(r\) and \(c\). When there is a given \(\omega'\), the interarrival time distribution can be defined as \(F_{\omega \rightarrow \omega'} (t|r,c)\), as demonstrated in (8). Therefore, \(\{X(t), t \geq 0\}\) is the CSMV process of employee turnover, with variables of \(r\) and \(c\).

Then, \(T_{\omega \rightarrow \omega} |r,c\) can be derived as (12), the multiplication of \(P\) and \(F\).

\[
T_{\omega \rightarrow \omega}(t|r,c) = \Pr(X_{\omega} = \omega', t_{\omega} < t \leq t + \omega_{\omega} | X_{\omega}, \omega, X_{\omega}, \omega, r, c). \quad (9)
\]

\[
P_{\omega \rightarrow \omega} |r,c\) = \Pr(X_{\omega} = \omega' | X_{\omega} = \omega, r,c) = \lim_{t \rightarrow t_{\omega}} T_{\omega \rightarrow \omega}(t|r,c). \quad (10)
\]

\[
F_{\omega \rightarrow \omega}(t|r,c) = \begin{cases} 
\frac{T_{\omega \rightarrow \omega}(t|r,c)}{P_{\omega \rightarrow \omega} |r,c\} & \text{if } P_{\omega \rightarrow \omega} |r,c\ > 0 \\
1 & \text{if } P_{\omega \rightarrow \omega} |r,c\ = 0 
\end{cases} \quad (11)
\]

\[
T_{\omega \rightarrow \omega}(t|r,c) = (P_{\omega \rightarrow \omega} |r,c\) F_{\omega \rightarrow \omega}(t|r,c). \quad (12)
\]

Assume that the state-transition cycle is one year, then \(\omega\) denotes job-state at the beginning of this year. Let \(d\) (\(d=0,1,2, \ldots\)) be the duration time of \(\omega\) at the beginning of this year, then the transfer probability to \(\omega'\) within this year can be derived as (13), with variables of \(r\) and \(c\). Let \(N_{d}(r)\) be the amount of employee whose duration time of \(\omega\) is exactly \(d\), then in (14), \(N_{d}'(r)\) denotes the amount of employee whose job-state is \(\omega\), in (15), the amount of employee whose job-state transfer to \(\omega'\) on condition of \(\omega\) is calculated by multiplying employee amount on \(\omega\) and

\[
T_{\omega \rightarrow \omega}(t|r,c) = \sum_{d=0}^{\infty} P_{\omega \rightarrow \omega} |r,c| F_{\omega \rightarrow \omega}(t|d,r,c). \quad (13)
\]

\[
N_{d}'(r) = N_{d}(r) \cdot P_{\omega \rightarrow \omega} |r,c| \cdot F_{\omega \rightarrow \omega}(t|d,r,c). \quad (14)
\]

\[
N_{\omega} (r) = \sum_{d=0}^{\infty} N_{d}'(r). \quad (15)
\]
transfer probability from $\omega$ to $\omega'$. Therefore, the total amount of employee whose job-state transfer to $\omega'$ can be calculated as (16), and $N^\omega_{i+1}^\omega'$ is the amount of employee on state $\omega'$ at the beginning of next year. In (17), it should be noted that $\in\Xi$ denotes an accessible process, e.g. in $\omega\in\Xi\omega'$, $\omega$ is the current job-state whose objective state is $\omega'$.

$$T_{\omega\to\omega'}(\tau,c) = \Pr[X_{\omega}\to\omega', d_{\omega}<t_{\omega}\to\omega'] - d_{\omega}\mathbb{P}[X_{\omega}\to\omega', \omega, r, c]$$

$$= T_{\omega\to\omega'}(d_{\omega}|r,c) - T_{\omega\to\omega'}(d_{\omega}|r,c).$$

$$N^\omega(\tau) = \sum_\omega N^\omega(\tau).$$

$$N^\omega(\tau) = \sum_\omega N^\omega(\tau)N^\omega_{i+1}^\omega'(\tau,c) = \sum_\omega N^\omega_{i+1}^\omega'(\tau,c).$$

$$N^{\omega\to\omega'}(\tau,c) = \sum_\omega N^{\omega\to\omega'}(\tau,c).$$

$$N^{\omega\to\omega'}(\tau) = \sum_\omega N^{\omega\to\omega'}(\tau).$$

$$N^{\omega\to\omega'}(\tau) = \sum_\omega N^{\omega\to\omega'}(\tau)\sum_\omega N^{\omega\to\omega'}(\tau).$$

These calculations have been specifically given in Fig.1, which displays the transition among groups as well.

**Figure 1.** State-transition based on employee’s transfer.

### 3.3 Variable adjustment

After state-transition in this year, the structure of groups will be changed next year, thus we propose variable adjustment to reset $\tau$ and $c$ at the beginning of next year.

As a natural factor of $\tau$, $\tau'$ is constant no matter the job-state transfers or not. On the contrary, $i$ will change as job-state transfers. In this section, turnover means promote, exchange or training instead of recruit, retire, resign or disposal; the reason is that recruit is an initialization of job-state and there is no objective state of retire, resign or disposal. Thus, both $\bar{\tau}$ and $\bar{\tau}'$ will be increased by $s$ at the beginning of next year if employee turnover occurs in this year. For $\bar{\tau}$, it will be zeroed if the turnover is promote; its value will be kept if the turnover is exchange or training for there is no influence on position grade of these two.

As for the environment variable, $c$ will be reset for each job-state group, which can be calculated according to the employee amount of each state-transition.

In summary, there are 5 steps of employee turnover prediction, as shown in Fig.2.

- **Step1:** Group the employee data according to job-state, construct hierarchical work of employee turnover, and distinguish state-transition among groups.
- **Step2:** Calculate environment variable for each group, count employee amount on condition of natural factor, and count employee amount of each group, based on statistics of employee data.
- **Step3:** Use state-transition function to predict the amount of employee transfers among groups on condition of adjoint variable.
- **Step4:** Calculate the employee amount, employee amount of each group and the employee amount on condition of adjoint variable, and update the environment variable for each group, based on current employee amount and the prediction of employee turnover.
- **Step5:** Repeat step 3 and 4 to predictive the employee structure of next year.
4 Case Study for A State-owned Enterprise

In order to demonstrate how state-transition and CSMK work in reality, here we provide a case study of Chinese state-owned enterprise. We select the employee data in recent 11 years, from 2006 to 2016, within nearly 6 million state-transition records. All properties \((i, j, k, \tau, c\) and \(d\)) mentioned in Section III can be acquired directly or calculated indirectly. Additionally, the true employee amount of each state group \((U_{\omega})\) has been involved in these dataset, denoted as \(N_{\omega}\).

Based on the employee turnover rules of this organization, job-state groups are classified as Fig.3, where \(I=8\) \((i=0,2, \ldots,7)\). Actually, there is a limitation that some turnover could only occur on several grades. To be specific, as in (18), retire and disposal only occurs when \(i=(6,7)\), recruit occurs only when \(i=(0,1,\ldots,4)\).

Subsequently, state-transition and CSMK are applied to predict employee turnover according to the 5 steps of prediction, then we can get the predicted employee amount of each job-state group. And the prediction result is demonstrated as Fig.4 and Fig.5.

\[
\begin{align*}
\Omega' &= \{(ijk)' | i \in I \} \\
\Omega'' &= \{(ijk)^{'} | i \in I \} \\
\Omega' &= \{(ijk)'' | i = 6,7 \} \\
\Omega' &= \{(ijk)^{''} | i = 6,7 \} \\
\Omega'' &= \{(ijk)^{'''} | i = 0,1,\ldots,4 \}.
\end{align*}
\]
Figure 3. Groups of job-state based on employee turnover rules.

Figure 4. Transfer probabilities among job position grades on condition of education degrees.
In Fig.4, we provide the transfer probabilities among different job grades (from 0 to 7), on condition of different education degrees. As for this four subfigures, degree equals to 0, 1, 2 and 3, from left to right and then from top to bottom, respectively. It apparent that the transfer always decrease slowly when \(i \leq 3\) or \(i \geq 5\), while it decrease rapidly when \(3 < i < 5\). In addition, when \(i \leq 5\), the higher the degree is, the higher the transfer probability. Specially, when \(e=0\), the employee’s education degree is lower than bachelor \((e=1)\). Thus we can conclude that transfer probabilities decrease as position increase, and the employ are always with higher promote probability on lower or higher position, while the employee on medium position transfers less frequently.

**Figure 5.** Comparison of predicted amount with the real amount of employee on condition of education degrees

In the line chart of Fig.5, we compare our predicted amount with the real amount of employee on condition of education degrees, from 2006 to 2016. The employee amount of each group (degree) in different years has been normalized according to the amount of 2016. It comes to conclusion that the amount of employee with higher education degree \((e \geq 1)\) increases over time, and the increase rate is positively relative to their degree. But the amount of employee with lower degrees decrease rapidly. Which suggests that the education level of the employee structure tends to be higher and higher. The average prediction accuracy is calculated as 0.855.

The case study proves the effect of our proposed method of turnover prediction.

5 Conclusion

There have been large amount studies on employee turnover, including its influence factors, modeling and prevent strategies. However, little attention has been paid to Chinese enterprise as well as the internal transfer within organization. Therefore, we propose the models of state-transition and CSMK to describe the turnover as well as to predict the turnover. These methods have been proved effective in reality by a case study of the employee records data, from 2006 to 2016, provided by a Chinese stated-owned enterprise.

Our study makes a primary contribution to predicting employee turnover of Chinese enterprise, both the inter-organization and the internal transfer. For managers, especially managers of Chinese enterprise, proposed method provides effective supports of developing recruitment plans, promote rules, and retire regulations. This turnover prediction method can be generalized to any organization that with similar structure and employee rules. In our prediction model, we didn’t consider the turnover demotion, which will be an interesting direction. And more influence factors should be considered in our further study, which might be individual, organizational or environmental.

Acknowledgment

This work is supported by the Natural Science Foundation of China (grant nos. 71522014, 71301165, 71690233, 71331008, 71671186 and 61473301).
References


