

## 대기행렬 모형을 사용한 기업 업무절차의 수행시간 예측\*

하병현<sup>1</sup>, 배준수<sup>2</sup>, 강석호<sup>1</sup>

<sup>1</sup>서울대학교 산업공학과/<sup>2</sup>전북대학교 산업정보시스템공학과

### 초 록

합리적인 업무 수행도의 예측을 통해 기업은 기존 업무절차의 평가뿐 아니라 업무 개선방안과 새로운 업무의 설계기준을 제시할 수 있다. 본 연구는 업무효율지표들 중 가장 중요한 요소인 업무절차의 수행시간을 예측하는 모형을 제시한다. 일반적으로 기업의 업무는 예측가능하며 장기적으로 안정된 성격을 가진다. 우리는 이러한 특성을 바탕으로 한 대기행렬 모형을 구축하고 그것을 분석하여 정적인 방식의 업무실행 시 수행시간을 예측하였다. 그리고 모형의 성능을 시뮬레이션 기법을 사용하여 평가하였다.

### 1. 서 론

고객중심의 기업환경에서 경쟁우위를 달성하기 위해서는 효율적인 업무절차의 관리가 필수적이다. BPMS(Business Process Management)는 업무절차를 관리하기 위한 지속적인 노력으로 탄생했다. 기존의 WFMS (Workflow Management Systems)는 내부의 업무절차를 개발, 실행, 관리하는 초석을 마련하였으며, 최근에는 이러한 결실을 기업 간 업무통합에 적용할 수 있도록 확장한 BPMS를 활용하고 있다[1].

BPMS는 작업을 참여자에게 적절하게 부여해 업무절차의 효율적인 수행을 제어한다. BPMS가 처리하는 업무절차는 미리 정의될 수 있고 수행 순서와 의미가 명확히 서술될 수 있어야 한다. 업무절차는 실행하기 전에 수행하는 작업들의 순서가 완전히 결정되는 것이 아니라 실행 도중에 업무의 진행 상황에 따라 작업들을 선택하는 특징이 있다. 또한 그 실행이 아주 오랫동안 진행이 되더라도 안정성 있고 예측 가능하다. 그리고 일반적으로 특정 작업을 수행하는 참여자의 자원은 한정되어 있다. 따라서 한꺼번에 많은 업무를 처리해야 하는 상황이 발생할 경우,

참여자의 수에 비해 수행해야 하는 작업들의 양이 증가하고, 일정한 작업을 처리할 수 있는 참여자는 역할에 의해 제한되어 있으므로, 다양한 작업을 효율적으로 처리하기 위해서는 정교한 업무할당규칙이 필요하다.

업무절차의 효율성 지표들 중 가장 중요한 요소는 수행시간(cycle time)이다. 업무 수행시간은 고객의 요구를 받은 시점부터 그와 관련된 모든 업무를 처리하여 고객에게 서비스를 제공하는 데까지 소요된 시간을 의미한다. 업무절차의 수행시간을 정확히 예측할 수 있다면 기업은 기존 업무절차의 평가뿐 아니라 업무 개선방안과 새로운 업무의 설계기준을 제시할 수 있다.

본 연구에서는 참여자의 자원 용량을 고려하여 업무절차의 수행시간을 예측하는 모형을 제시한다. 먼저, 참여자를 대기행렬 시스템으로 모형화하여 개개 작업의 수행시간을 예측한 후, 여러 작업을 묶는 가상작업의 개념을 도입하여 전체 업무절차의 효율성을 구한다.

### 2. 관련 연구

전통적으로 시스템의 수행시간 단축에 관한 문제는 생산 분야에서 가장 많이 연구되어왔다. 생산 공정의 일정계획문제는 수행시간과 관련된 대표적인 문제이다. 하지만, 일정계획문제의 해법을 업무절차의 분석에 도입하는 데는 특유의 복잡성과 확률적인 성격에 의해 쉽지 않다. 프로세스의 효율적인 실행을 다루는 또 하나의 전통적인 분야는 프로젝트 관리 분야이며, 주공정을 분석함으로써 프로젝트의 진행을 제어한다. 이 분야의 대표적인 방법인 PERT와 CPM은 보통 자원제약이 없다는 가정에서 주공정을 분석하거나, 또는 주공정을 개선하기 위한 자원 능력을 계획하는데 사용된다. 따라서 본 연구에서 다루는 업무를 수행하는 자원이 제약되어있는 경우에 적용하기는 한계를 가진다.

대기행렬 모형을 사용하여 업무절차의 분석은 다양한 방법과 다양한 목적을 위해 이루어져 왔다.

\* 본 연구는 과학기술부의 21C프론티어사업, 과제번호 (04A2-B1-31)의 지원으로 수행되었음.

먼저 Narahari et al.은 신제품 개발과정에서 조직의 부서를 대기행렬시스템으로 모형화 해서 수행시간을 분석하고 그것을 줄이는 여러 방법을 제시하였다[7]. 업무절차를 구성하는 작업을 대기행렬시스템으로 모형화 해서 업무의 수행 시나리오를 지키기 위한 용량을 계획하거나[10], 주공정을 구하는 것도 연구되었다[4]. 업무절차는 페트리넷을 사용해 설계를 할 수 있다. Stochastic Workflow Net (SWN)은 페트리넷을 사용해 설계한 업무절차를 대기행렬을 통해 분석하는 모형이다. SWN을 사용하여 업무수행도의 분석과 용량계획이 가능하다[11].

우리는 이미 업무절차의 수행 시 참여자의 부하 측정을 대기행렬 모형을 사용하여 예측하는 방식을 제안하였다[2]. 이 연구는 참여자의 용량을 고려했다는 점에서 선행 연구들과 차별성을 가진다. 하지만, 이 모형을 확장하여 업무절차의 수행시간을 예측하기 위해서는 병렬적으로 수행되는 작업들의 실행시간을 예측하여야 한다. 대기행렬 네트워크에서 이러한 병렬수행 작업의 집단은 Fork-Join 대기시스템으로 모형화된다. 간단한 경우의 Fork-Join 대기시스템에 대해 많은 연구자들이 정확한 해를 구하거나 근사해와 그 해의 정확성에 관해 연구하였다[6][3]. 하지만 이와 같은 접근 방법은 표 등을 이용한 단계적인 해석 방식을 사용하여 계산량이 많을 뿐만 아니라 업무절차와 같은 복잡한 구조를 분석하는데 한계를 가진다. Rajaraman & Morgan은 일반화된 지수분포를 사용하여 정확성을 희생하는 대신 보다 간편한 방법을 제시하였다[8]. 본 연구에서는 이 방법을 사용하여 업무의 병렬처리 시 작업시간을 예측한다.

### 3. 개별 작업의 실행시간 예측

본 연구에서는 업무절차의 수행시간을 예측하기 위해 분석적절차모형(Analytic Process Model)을 사용한다[5]. 이 모형이 제공하는 정보들을 바탕으로 정적 업무수행 방식을 사용하여 업무절차 수행 시 개별 작업의 수행시간을 대기행렬 모형을 사용하여 예측한다.

#### 3.1 분석적절차모형

분석적절차모형과 관련 정의들은 다음과 같다. 정의 1 (분석적절차모형). 분석적절차모형은  $(\Phi, T, L, A, R)$ 으로 정의되며 다음과 같은 특성을 지닌다[5].

i)  $\Phi$ 는 고객 도착률이다.

ii)  $T$ 는 작업의 집합이다.

iii)  $L \subseteq T \times T$ 는 링크의 집합이며,  $e_l$ 은  $l \in L$ 의 구조적수행확률이다.

iv)  $A$ 는 참여자의 집합이다.

v)  $R \subseteq T \times A$ 은 작업을 처리하는 참여자의 책임을 나타내며,  $(t, a) \in R$ 를 만족하는  $\mu_{t,a}$ 는 참여자  $a$ 가 작업  $t$ 를 처리하는 평균 서비스률이다.

정의 2 (예상실행비율).  $t$ 의 예상실행비율  $f_t$ 는 업무절차가 한 번 수행될 때 작업  $t$ 가 실행되는 회수이다[5].

정의 3 (작업도착률). 작업  $t$ 의 작업도착률  $\lambda_t$ 는 업무절차가 고객에 의해 지속적으로 수행될 때 그 작업이 실행되는 비율이다[5].

#### 3.2 정적 업무수행 규칙

본 연구에서 분석하는 작업할당 규칙은 정적 업무수행 규칙이다. 이 규칙을 따르면 작업자는 자신만의 워크리스트를 사용하고, 워크리스트에 있는 작업들 중 먼저 할당된 작업을 먼저 처리하는 선입선출 방식을 사용하여 일을 수행한다. 이 방식은 일견 공유 워크리스트를 사용하고 수행할 작업을 동적으로 선택하는 것보다 수행도가 낮을 것으로 예측되지만 EDD 등의 다른 방식보다 오히려 좋은 수행도를 보여준다는 것이 알려져 있다. (Ha et al, 2004) 또한 이 규칙을 사용할 경우 단순한 실행 형태로 인해 분석 가능한 대기행렬 네트워크를 구축할 수 있다. 정적 업무수행 규칙을 사용하기 위해 특정 작업을 수행할 시점에서 어떤 참여자에게 그 작업을 할당할 확률을 사용한다.

정의 4 (작업할당확률).  $(t, a) \in R$ 인 작업  $t$ 를 참여자  $a$ 에게 할당하는 작업할당확률  $p_{t,a}$ 는 업무절차가 수행될 때 그 작업을 수행하기 위해 그 참여자가 선택될 확률이다[5].

#### 3.4 참여자의 대기모형

업무절차를 수행하는 중 참여자가 수행하는 작업은 작업할당확률에 따라 결정된다. 그리고 참여자에게 주어지는 작업은 그 참여자의 능력에 따라 각각 다른 서비스률로 처리된다. 따라서 업무절차를 수행하는 참여자는 다계층 고객들을 처리하는 대기행렬시스템으로 모형화 할 수 있다. 그리고 전체 업무절차는 참여자의 대기행렬시스템들이 연결된 대기행렬 네트워크로 형성된다. 본 연구에서는 계산의 편의성을 위해 참여자로 도착하는 작업은 포아송 분포를 따른다고 가정하였고 참여자는 작업의 도착순서에

관계 없이 확률적으로 작업의 종류를 결정하여 처리한다고 가정하였다. 이와 같은 가정을 사용하면 참여자는 M/H/1 대기행렬시스템으로 모형화 할 수 있으며, 본 연구에서 그것의 성능지표를 사용하여 업무절차의 성능지표를 구한다. 이와 같은 모형을 사용할 경우의 부정확성은 평가에서 논의하였다.

이와 같은 경우 참여자  $a$ 에게 주어지는 작업  $t$ 의 도착률( $\lambda_{t,a}$ )과 모든 작업의 총 도착률( $\lambda_a$ ), 업무부하( $\rho_a$ ), 서비스시간의 이차, 삼차 적률( $E(S_a^2)$ ,  $E(S_a^3)$ ), 작업 실행시간의 일차, 이차 적률( $E(T_a)$ ,  $E(T_a^2)$ )은 다음과 같이 정의된다.

$$\lambda_{t,a} = p_{t,a} \lambda_t, \quad (t,a) \in R \quad (1)$$

$$\lambda_a = \sum_{t,(t,a) \in R} \lambda_{t,a} \quad (2)$$

$$\rho_a = \sum_{t,(t,a) \in R} \frac{\lambda_{t,a}}{\lambda_a} \frac{1}{\mu_{t,a}} \quad (3)$$

$$E(S_a^2) = \sum_{t,(t,a) \in R} \frac{\lambda_{t,a}}{\lambda_a} \frac{2}{\mu_{t,a}^2} \quad (4)$$

$$E(S_a^3) = \sum_{t,(t,a) \in R} \frac{\lambda_{t,a}}{\lambda_a} \frac{6}{\mu_{t,a}^3} \quad (5)$$

$$E(T_a) = \frac{\lambda_a E(S_a^2)}{2(1-\rho_a)} + \frac{\rho_a}{\lambda_a} \quad (6)$$

$$E(T_a^2) = \frac{\lambda_a E(S_a^3)}{3(1-\rho_a)} + \frac{\lambda_a^2 E(S_a^2)^2}{2(1-\rho_a)^2} + \frac{E(S_a^2)}{(1-\rho_a)} \quad (7)$$

### 3.5 개별 작업의 실행시간

참여자의 응답시간을 위와 같이 예측하였을 때 각각의 참여자들의 작업실행 시간은 서로 독립이므로 특정 작업  $t$ 의 수행시간( $T$ )과 일차, 이차 적률( $E(T)$ ,  $E(T^2)$ )은 다음과 같다.

$$T_t = \sum_{a,(t,a) \in R} \frac{\lambda_{t,a}}{\lambda_a} T_a \quad (8)$$

$$E(T_t) = \sum_{a,(t,a) \in R} \frac{\lambda_{t,a}}{\lambda_a} E(T_a) \quad (9)$$

$$E(T_t^2) = \sum_{a,(t,a) \in R} \frac{\lambda_{t,a}^2}{\lambda_a^2} E(T_a^2) + \sum_{\substack{a_1,(t,a_1) \in R, a_2,(t,a_2) \in R, \\ a_1 \neq a_2}} \frac{\lambda_{t,a_1} \lambda_{t,a_2}}{\lambda_{a_1} \lambda_{a_2}} E(T_{a_1}) E(T_{a_2}) \quad (10)$$

## 4. 업무절차의 수행시간 예측

3장에서 예측한 개별 작업의 실행시간을 예측한 결과를 사용하면 전체업무의 수행시간을 예측할 수 있다. 우리는 전체업무의 수행시간을 구하기 위해 일반화된 지수분포를 사용하였다.

작업들 중 가장 안쪽에 위치한 작업들로부터 주변 작업들을 합쳐, 그 결과를 합쳐진 작업을 대표하는 새로운 가상작업을 생성해 나가는 방식을 사용해, 최종적으로 업무절차를 하나의 가상작업으로 보고 그것의 수행시간을 구한다. 작업들을 합쳐 새로운 가상작업으로 만들 때 새로운 작업의 정확한 분포를 구하기 힘들다. 따라서 이 계산을 단순화시키기 위해 합쳐지는 작업의 일차와 이차 모멘트만 사용하여 새로운 작업을 생성한다. 따라서 가상작업은 일반화된 지수분포[8]를 따른다.

### 4.1 순차 작업의 처리

순차 가상작업  $t_1, t_2, \dots, t_n$ 이 합쳐져 만들어지는 새로운 가상작업  $t_s$ 의 일차, 이차 적률( $E(T_s)$ ,  $E(T_s^2)$ )은 다음과 같다.

$$E(T_s) = \sum_i E(T_i) \quad (11)$$

$$E(T_s^2) = \sum_i E(T_i^2) + \sum_{i,j,i \neq j} 2E(T_i)E(T_j) \quad (12)$$

### 4.2 루프 구조의 처리

하나의 가상작업  $t$ 가 루프 구조에 의해  $p$ 의 확률로 재실행되는 경우 그것을 대표하는 가상작업  $t_t$ 의 일차, 이차 적률( $E(T_t)$ ,  $E(T_t^2)$ )은 다음과 같다.

$$E(T_t) = \frac{1}{1-p} E(T) \quad (13)$$

$$E(T_t^2) = \frac{1}{(1-p)^2} E(T^2) \quad (14)$$

### 4.3 병렬 구조의 처리

두 개의 가상작업  $t_1, t_2$ 가 병렬로 처리되어야 하고 각각이 처리될 구조적수행확률이  $l_1, l_2$ 라고 하면,  $t_1$ 만 실행될 확률( $p_1$ ),  $t_2$ 만 실행될 확률( $p_2$ ),  $t_1$ 과  $t_2$ 가 동시에 실행될 확률( $p_{12}$ )은 각각 다음과 같다.

$$p_1 = 1 - l_2 \quad (15)$$

$$p_2 = 1 - l_1 \quad (16)$$

$$p_{12} = l_1 + l_2 - 1 \quad (17)$$

가상작업  $t_1$ 과  $t_2$ 의 실행시간인  $T_1, T_2$ 는 일반화된 지수분포이다. 따라서 이것을 이용하여  $m_1, m_2, d_1, d_2$ 를 정의한 후, 두 작업이 동시에 실행될 때의 실행시간의 일차, 이차 적률( $E(T_{12})$ ,  $E(T_{12}^2)$ )을 구하면 다음과 같다.

$$m_1 = \sqrt{E(T_1^2) - E(T_1)^2} \quad (18)$$

$$d_1 = E(T_1) - m_1 \quad (19)$$

$$m_2 = \sqrt{E(T_2^2) - E(T_2)^2} \quad (20)$$

$$d_2 = E(T_2) - m_2 \quad (21)$$

$$E(T_{12}) = m_1 e^{\frac{d_1}{m_1}} + m_2 e^{\frac{d_2}{m_2}} - \frac{m_1 m_2}{m_1 + m_2} e^{\frac{d_1 + d_2}{m_1 m_2}} \quad (22)$$

$$E(T_p^2) = 2m_1^2 e^{\frac{d_1}{m_1}} + 2m_2^2 e^{\frac{d_2}{m_2}} - 2 \frac{m_1^2 m_2^2}{(m_1 + m_2)^2} e^{\frac{d_1 + d_2}{m_1 m_2}} \quad (23)$$

위의 계산은 T1과 T2, 즉, 병렬적으로 수행되는 두 가상작업의 수행시간이 독립이라는 가정을 두고 이루어졌다. 여기에 대한 분석도 뒤에서 논의될 것이다.

위의 결과를 기반으로 병렬 구조를 나타내는 가상작업의 일차, 이차 적률( $E(T_p)$ ,  $E(T_p^2)$ )을 구하면 다음과 같다.

$$E(T_p) = p_1 E(T_1) + p_2 E(T_2) + p_{12} E(T_{12}) \quad (24)$$

$$E(T_p^2) = p_1^2 E(T_1^2) + p_2^2 E(T_2^2) + p_{12}^2 E(T_{12}^2) + p_1 p_2 E(T_1) E(T_2) + p_2 p_{12} E(T_2) E(T_{12}) + p_{12} p_1 E(T_{12}) E(T_1) \quad (25)$$

#### 4.4 업무절차의 수행시간

업무절차의 예상 수행시간(CT)은 최종적으로 만들어진 가상작업( $T_f$ )의 일차 적률이다.

$$CT = E(T_f) \quad (26)$$

### 5. 실험 및 평가

간단한 업무절차와 복잡한 업무절차 두 종류에 대하여 예측값과 시뮬레이션 결과를 비교하였다(그림 1).

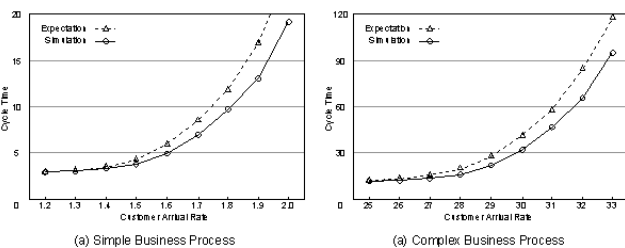


그림 1. 수행시간 예측값과 시뮬레이션 결과 비교

수행시간의 예측값이 시뮬레이션 결과보다 약 20~30% 크게 나오는 것을 알 수 있다. 그 이유는 크게 세 가지로 추측할 수 있다. 먼저 일반화된 지수분포의 사용으로 고차 적률이 고려되지 않으므로 가상 작업으로 합쳐갈수록 오류가 누적되는 것으로 파악된다. 다음으로 본 연구에서는 참여자를  $M/H, 1$ 으로 모형화 하였는데, 대기행렬 네트워크에서 입력은 종속성으로 인해 포아송 분포를 따르지 않는다는 것이 알려져있다[1]. 마지막으로 병렬처리 작업으로의 입력 역시 독립이 아니다.

### 6. 결론 및 추후과제

우리는 본 연구를 통해 참여자의 자원이 제한되어있는 상황에서 업무절차의 수행시간을 예측하는 알고리즘을 제시하고 그것의 성능을 검증하였다.

앞으로 보다 정확하고, 분석뿐 아니라 계획에도 사용할 수 있는 모형을 개발하는 것이 필요할 것으로 생각된다.

### 참고문헌

- [1] 이호우: 대기행렬이론. 도서출판 기술 (1996)
- [2] 하병현, 설현주, 배준수, 박용태, 강석호: 통계적 모형의 업무부하 균일화를 통한 비즈니스 프로세스의 효율화. IE Interface 16 (2003)
- [3] Baccelli, F., Massey, W.A., Towsley, D.: Acyclic Fork-Join Queueing Networks. Journal of the ACM 36 (1989) 615-642
- [4] Chang, D., Son, J., and Kim, M.: Critical Path Identification in the Context of a Workflow. Information and Software Technology 44 (2002) 405-417
- [5] Ha, B.-H., Bae, J., Kang, S.-H.: Workload Balancing on Agents for Business Process Efficiency based on Stochastic Model. In: Proceeding of 2nd International Conference on Business Process Management. Potsdam (2004)
- [6] Kim, C., Agarwala, A.K.: Analysis of the Fork-Join Queue. IEEE Transactions on Computers 38 (1989) 250-255
- [7] Narahari, Y., Viswanadham, N.: Lead Time Modeling and Acceleration of Product Design and Development. IEEE Transaction on Robotics and Automation 15 (1999) 882-896
- [8] Rajaraman, B., Morgan, T.W.: Approximate Analysis of the Average Delay in Parallel Program Execution. In: Proceeding of the Twenty-Sixth Hawaii International Conference on System Sciences, Hawaii (1993) 584-593
- [9] Smith, H., Fingar, P.: Business Process Management: The Third Wave. Meghan-Kiffer, Tampa (2003)
- [10] Son, J., Kim, M.: Improving the Performance of Time-Constrained Workflow Processing, Journal of Systems and Software 58 (2001) 211-219
- [11] Zerguini, L., van Hee, K.M.: A New Reduction Method for the Analysis of Large Work-flow Models. In: Proceedings of the Joint Annual Conference of the GI Special Interest Groups "Petrietze und verwandte Systemmodelle" and EMISA (Promise 2002), Potsdam (2002) 188-201