

The Prefetching Method in Mobile Environments

Jin-Ah Yoo¹⁾ · Tea-Wook Koh²⁾

Abstract

This paper proposes a mobile computing prefetching method providing the effective information about location change of mobile user or mobile computing in mobile information services. For mobile computing environments, there exist restrictions as like low bandwidth, latency and traffic. To solve those problems, a variety of techniques have been developed including caching and prefetching. In this paper we present a Statistical Poisson Prefetching Scheme using the reference count to provide a mobile user information that will be likely referenced in the near future. Comparing to existing methods in numerical results, the proposed method improves the prefetching performance to give the maximum effectiveness and reduces the failure rate of information searching.

Keywords : 위치인식(location-aware), 캐싱(caching), 포아송(poisson), 프리페칭(prefetching)

1. 서 론

최근 네트워크 기술이 급속도로 발전함으로써 컴퓨터의 크기가 휴대하기에 충분히 작아지면서 휴대성과 이동성이 가능한 모바일(Mobile)로 중심축이 바뀌고 있다. 무선 매체를 사용하는 영역 또한 크게 확대됨에 따라 무선인터넷 서비스는 위치기반의 기술등의 하드웨어적인 개발보다는 사용자 위주의 정보제공서비스에 주력하고 있다. 즉 무선통신 기술의 발전과 이동 가능한 통신 기기의 대중화로 인하여 사용자들은 위치나 장소에 관계없이 현재 상황에 적절한 정보를 제공받기를 원한다. 이러한 정보는 단순한 텍스트 수준을 넘어서 멀티미디어형식을 갖춘 종합적이고 고차원적인 정보서비스의 요구가 점차적으로 높아지고 있다(Tomasz, 1996). 이런 고차원적인 서비스는 사용자가 이동 중에 제공되어야함으로 환경의 변화와 관련된 컴퓨터와 요소등이 서로

1) 전북 전주시 완산구 효자동 3가 1200번지 전주대학교 통계학과
E-mail : gina@jj.ac.kr

2) 전북 전주시 완산구 효자동 3가 1200번지 전주대학교 통계학과 교수
E-mail : twkoh@jj.ac.kr

영향을 주게 된다. 무선 고유의 특성에 따른 문제점들로 인해 기존의 유선 환경에서 적용되던 기술들이 직접 적용되기에는 많은 제약을 가지고 있다.

이동환경은 낮은 대역폭, 높은 지연과 트래픽 그리고 잦은 연결의 재설정등은 커다란 장애요소로 남아 있다(B. Bruegge, 1996). 대역폭을 증가시키는 방법은 많은 비용 추가라는 단점이 있기 때문에 대역폭을 최대한 활용하는 방법을 찾는 것이 바람직하다. 이에 대한 대안으로 흔히들 캐싱기법을 사용하며, 크게 캐시교체부분과 프리페칭(prefetching)으로 구분되어 진다. 캐시교체는 이동 호스트에서 새로운 데이터를 업데이트하기 위하여 기존의 데이터 중에서 삭제할 데이터를 선정하는 것을 말한다. 프리페칭은 향후 사용 가능성이 높은 데이터를 미리 캐시에 가져다 놓는 것으로 이동 컴퓨팅 환경의 특성을 고려한다면 이의 역할은 중요하다 하겠다. 이렇게 필요한 데이터를 미리 프리페칭 함으로써 서버 측면에서는 무선 네트워크를 통해 전송하기 위하여 미리 필요한 데이터를 변화시킬 시간을 확보할 수 있고, 이동 호스트 측면에서는 다운로드 비용이 업로드에 비해 낮은 비용을 줄일수 있으므로 효과적이다(Dhananjay, 2002). 어떤 데이터를 프리페칭할 것인가를 결정하는 것도 중요하지만, 소형 이동 단말기를 사용하는 사람에게는 전력소모의 문제점이 있기 때문에 몇 개의 데이터를 프리페칭할 것인지를 결정하는 것도 중요하다 하겠다(Yin, 2002). 이러한 프리페칭 수에 제한을 두어 메모리를 효율적으로 관리하는 전략으로 속도기반의 프리페칭과 방문지의 이동 빈도수를 적용한 프리페칭을 들 수 있다.

지금까지 프리페칭에 관한 연구들 살펴보면 확률분포를 이용하여 적용한 사례를 찾아볼 수가 없다. 현실에서 프리페칭을 적용하는데 있어 분포를 이용하는 것은 중요한 문제이다. 따라서 본 논문에서는 모바일 사용자가 단위시간동안에 어떤 특정 셀을 이동하기 때문에 이동패턴을 근거로 방문하는 빈도를 이용하여, 통계적 방법의 하나인 포아송분포를 이용한 프리페칭 알고리즘을 구현하여 본다. 본 논문은 2장에서 기존의 속도기반의 프리페칭 방법들에 대한 기존연구를 살펴보고, 3장에서는 단말의 이동 패턴과 포아송 분포를 이용한 프리페칭 알고리즘을 제안한다. 4장에서는 제안된 프리페칭 방법과 기존의 방법들을 비교한다. 마지막으로 5장에서는 결론 및 추후 과제에 대해 제시한다.

2. 관련연구

모바일환경에서 사용자의 현재 상황을 충분히 반영할 수 있는 정보서비스 방법론은 중요한 기술요소로 인식되고 있고 활발한 연구가 되어지고 있다. 이러한 점들을 고려하여 지금까지 많은 프리페칭 방법들이 제시되었다.

분산 환경에서 사용자의 이동에 대한 로그 파일을 이용하여 사용자의 움직임과 접근되는 객체 속성을 고려한 패턴을 마이닝하고, 그 결과를 각 서버에 대한 데이터 할당(즉 서버측 프리페칭)에 이용하여 성능을 높이는 결과를 보였다(Peng, 2000).

반면, 위치 기반 서비스를 위한 이동 호스트 측면에서의 프리페칭 연구가 수행되었다. 즉 사용자가 미래에 필요로 할 정보를 예측하여 실제로 사용되기 전에 이동 호스트로 미리 이동시켜 놓음으로써 캐시 성능을 향상시켰다(Kubach, 2001).

이때 어떤 데이터를 프리페칭할 것인가를 결정하는 것도 중요하지만, 소형 이동 단말기를 사용하는 사람에게는 전력소모의 문제점이 있기 때문에 몇 개의 데이터를 프

리페칭할 것인지를 결정하는 것도 중요하다 하겠다. 이러한 프리페칭 수에 제한을 두어 메모리를 효율적으로 관리하는 전략으로 속도기반의 프리페칭과 방문지의 이동 빈도수를 적용한 프리페칭을 들 수 있다.

속도기반의 빈도수(김문자,2002)는 사용자의 위치는 이동속도와 이동방향에 의해서 표현된다는 가정을 이용하였다. 이동 사용자는 주어진 관측 시간간격동안 일정한 속도로 2차원 공간을 이동한다고 가정한다. 이때 이동 사용자의 이동거리와 이동방향을 정형화하기 위해서 이동속도 V 는 벡터 $\langle V_x, V_y \rangle$ 로 정의한다. 이때 V_x 와 V_y 는 정수형이며, 이들의 값은 각각 x 와 y 좌표에 대한 속도를 나타낸다. 사각의 프리페칭의 영역을 산출하기 위하여, 길이 d 를 산출하고, 상위정수값을 취한다. 이때, $|V_x| > |V_y|$ 라면 d 는 x 축이 되며, $|V_x| < |V_y|$ 라면 d 는 y 축에 놓는다. 또한 $|V_x| = |V_y|$ 이면 d 는 x 나 y 축에 무관하다. 또한 이동 속성을 유지하기 위해서 V_x 와 V_y 값 중의 한 벡터 값이 0일 경우, 실제 속력 값을 반영하기 위해서 0인 벡터 값을 0이 아닌 속력 값으로 대체한다. 사각형영역을 다른 한 면의 폭 $w = \text{ROUNDUP}\left(\frac{|V_x| + |V_y|}{2}\right)$ 로 산출하며, 상위정수값을 취한다. 이때, $|V_x| > |V_y|$ 라면 d 는 x 축이 되며 w 는 y 축이 되어지고, $|V_x| < |V_y|$ 라면 d 는 y 축이 되며 w 는 x 축이 된다. 또한 $|V_x| = |V_y|$ 이면 w 는 x 나 y 축에 무관하다. 또한, 이동방향은 단순히 기울기 정보를 고려하므로 속도 값들의 부호는 무시된다. 그러나 길이 d 에서와는 다르게 V_x 와 V_y 벡터 값 중 어느 하나가 0일 경우, 이는 기울기를 내포하지 않으므로 0을 그대로 사용한다. 이러한 사각형 프리페칭 영역 설정 전략은 이동속력과 이동방향에 대한 파라미터 값이 유연하게 사각형 프리페칭 영역 구조에 수용되어 사용자의 이동 경로 상에 해당하는 위치분할을 포함함은 물론 미래에 유효하게 활용될 것으로 예상되는 주변의 위치분할들을 포함하는 결과를 갖는다. 만약 단말의 이동 속도가 $V = \langle -3, 3 \rangle$ 인 경우로 프리페칭영역의 크기는 길이와 폭이 각각 5와 3으로 산출되어 가상적인 프리페칭영역이 설정된다. <그림1>에서 속도기반의 프리페칭의 수는 15개 셀이다.



<그림 1> 속도기반의 프리페칭

빈도수 기반의 프리페칭(Park,2004)은 속도기반의 프리페칭을 기본으로 사용자의 이동패턴에 따른 공간 지역성의 방문 횟수를 이용하였다. <그림2>에서 속도기반의 프리페칭의 수는 15개 셀이고, 이것의 평균 빈도수는 11.3이다. 빈도수 기반의 프리페칭에서는 평균빈도수 이상인 셀들만을 선택한 결과 9개의 프리페칭을 얻는다.

3	16	7	8	13	15	10	9
17	14	5	3	16	7	4	13
20	9	20	11	10	16	12	5
14	15	12	6	14	9	2	14
12	6	12	17	19	16	13	8
10	9	9	8	4	11	18	5
14	5	17	2	11	6	30	20

<그림 2> 빈도수기반의 프리패칭

3. 제안한 프리패칭 알고리즘

이산확률분포 중에서 이항분포만큼 중요한 위치를 차지하는 분포가 포아송분포이다. 우리가 흔히 “특정한 현상이 일어나는 횟수”에 대한 모형으로 포아송분포를 이용한다. 그 동안 포아송분포에 대한 모형연구는 상당히 많은 부분에서 이루어져 왔다. 모수 형태가 개수나 비율에 대해서 모형 선택에 따라 추정방법 역시 다양하다.

포아송분포는 많은 실제 현상을 모형화한다. 포아송 확률변수의 값이 음이 아닌 정수이므로 관심 있는 종류의 수에 대한 확률적 현상을 포아송 분포를 가정하여 모형화하는데 사용할 수 있다. 임의의 셀을 방문한 포아송 분포를 이용한 확률밀도함수는 다음과 같다.

$$P(k) = \frac{e^{-\mu} \mu^k}{k!} \quad k = 0, 1, 2, \dots$$

여기서 k 는 임의의 셀이고, μ 는 빈도수 기반의 프리패칭에서 이동영역의 전체 평균 값을 의미한다.

포아송 분포를 고려하면 다음이 성립되어야 한다.

- ① $\mu > k$ 이면, $\frac{\mu^{(k-1)} e^{-\mu}}{(k-1)!} < \frac{\mu^k e^{-\mu}}{(k)!}$
- ② $\mu < k$ 이면, $\frac{\mu^{(k-1)} e^{-\mu}}{(k-1)!} > \frac{\mu^k e^{-\mu}}{(k)!}$
- ③ μ 가 정수이고 $\mu = k$ 이면, $\frac{\mu^{(k-1)} e^{-\mu}}{(k-1)!} = \frac{\mu^k e^{-\mu}}{(k)!}$

그 이유는 포아송 밀도함수에서

$$\frac{P(X=k-1)}{P(X=k)} = \frac{\mu^{(k-1)} e^{-\mu} / (k-1)!}{\mu^k e^{-\mu} / (k)!} = \frac{k}{\mu}$$

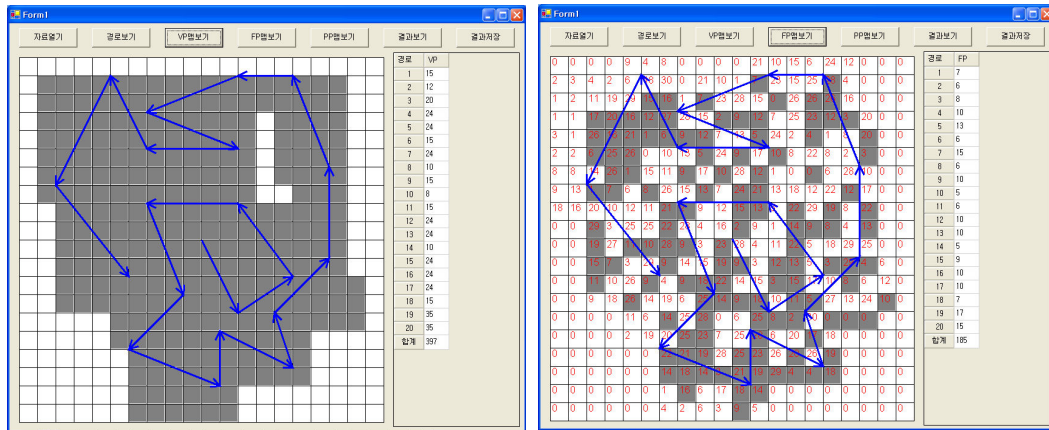
이므로, $\mu > k$ 이면 1보다 작고 $\mu < k$ 이면 1보다 크고 $\mu = k$ 이면 1임을 알 수 있기 때문이다.

포아송 기반의 프리패칭의 결과 속도기반의 프리패칭의 영역(15개)에 빈도수 프리패칭의 빈도(각 셀의 빈도 k 값)를 이용하여 포아송 분포를 산출한다. 그 결과

이동 사용자는 일정한 시간간격 동안 일정한 속도와 방향을 가지고 2차원 공간으로 구성된 이동 가능한 위치영역 내에서 이동한다. 시뮬레이션의 과정동안 이동 사용자는 좌표(10, 10)의 위치 영역을 시작점으로 설정한다. 시작점을 기점으로 (2, -4)→(3, 2)→(-3, 4)→(-5, 0)→(2, -5)→(-3, -3)→(5, -2)→(0, 3)→(4, -2)→(-1, 3)→(3, 3)→(0, 5)→(-2, 5)→(-3, 0)→(-5, -2)→(5, -2)→(-5, 0)→(-2, 4)→(-3, -6)→(4, -5) 순으로, 미리 주어진 좌표 값에 따라 다음 목적지의 위치영역까지 이동하는 과정을 20번 반복하는 것으로 하여 속도기반의 가상적인 영역을 설정하였다. 또한 가상적인 프리패칭 영역은 이동사용자의 현재 위치에서의 특정 속도를 가정하여 프리패칭 영역을 설정하는 것으로 하였다.

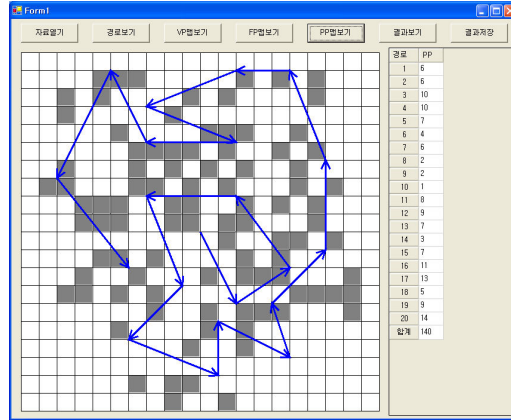
4.1 프리패칭 수

먼저 시뮬레이션 단말 이동 시나리오에서 제시한 이동순서를 기준으로 단말에 적재되는 정보분할의 수를 비교함으로써 제안된 프리패칭 전략의 유효성을 검토하였다. 각각의 경우에 시뮬레이션의 실행환경을 동일하게 설정하였다. 속도기반에 의한 프리패칭 전략은 속도기반 이동모델에 근거하여 사용자의 이동속도와 이동방향을 활용하여 프리패칭 영역의 길이와 폭을 설정함으로써 속도가 빠르면 그만큼 프리패칭할 영역의 길이와 폭이 길어지고 속도가 느리면 짧아지게 되었다. 하지만 편의상 본 시뮬레이션의 이동 시나리오에서는 이동순서에 따른 단말이 특정 속도를 가진다고 가정함으로써 특정 속도에 따른 프리패칭 영역을 설정하는 것으로 하였다. 그 결과는 <표 3> 와 같다. 속도기반의 프리패칭은 268개의 정보분할을 프리패칭하고 있으며 빈도수기반의 프리패칭은 129개로 나타났다.



<그림 5> 속도기반의 프리패칭 시뮬레이터 <그림 6> 빈도수기반의 프리패칭 시뮬레이터

이에 반하여, 제안된 포아송분포를 적용한 프리패칭 전략은 기존의 제안 알고리즘보다 프리패칭되는 정보분할의 수를 줄일 수 있게 되었다. 즉 사용자의 이동 동작의 특성을 프리패칭영역의 제한에 효율적으로 적용하고 있으므로 <그림5,6,7>에서 보이는 바와 같이 속도기반의 프리패칭전략과 비교하여 프리패칭되는 정보분할의 수는 많

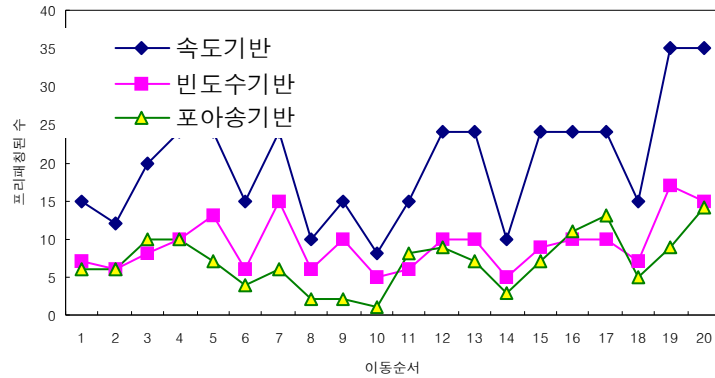


<그림 7> 포아송기반의 프리패칭 시물레이터

이 줄어들었음을 볼 수 있었다. 사용자의 이동에 따라 95개의 정보분할을 프리패칭하고 있으며 이동시마다 평균 7.0개의 정보분할을 프리패칭하는 것으로 산출되었다. 즉 사용자의 이동 경로상에 위치한 활용 가능성이 높은 정보분할만을 프리패칭함으로써 속도기반의 프리패칭에 비해서 35%의 향상을 보이고 있다.

<표 1> 프리패칭수

이동경로	x	y	d	w	속도기반	빈도수기반	포아송기반
1	2	-4	3	5	15	7	6
2	3	2	4	3	12	6	6
3	-3	4	4	5	20	8	10
4	-5	0	8	3	24	10	10
5	2	-5	4	6	24	13	7
6	-3	-3	5	3	15	6	4
7	5	-2	6	4	24	15	6
8	0	3	2	5	10	6	2
9	4	-2	5	3	15	10	2
10	-1	3	2	4	8	5	1
11	3	3	5	3	15	6	8
12	0	5	3	8	24	10	9
13	-2	5	4	6	24	10	7
14	-3	0	5	2	10	5	3
15	-5	-2	6	4	24	9	7
16	5	-2	6	4	24	10	11
17	-5	0	8	3	24	10	13
18	-2	4	3	5	15	7	5
19	-3	-6	5	7	35	17	9
20	4	-5	5	7	35	15	14
정보분할수					268	129	95



<그림 8> 단말기에 프리페칭 수

4.2 정보검색 실패율

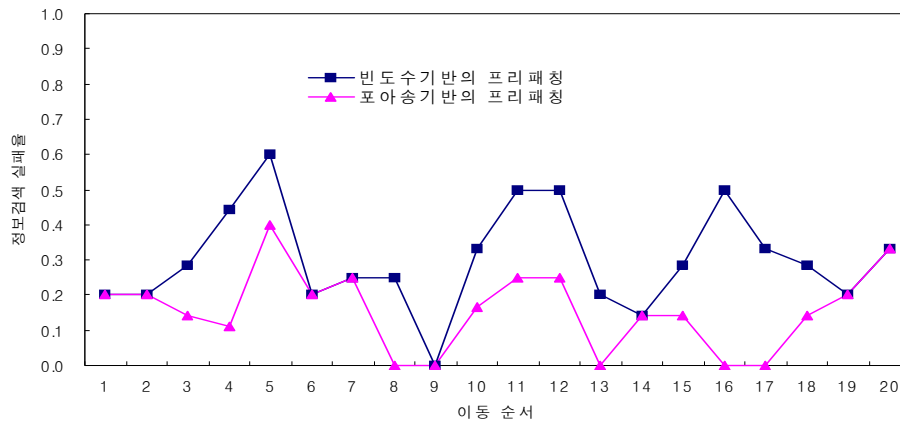
지금까지 프리페칭 정보분할의 수의 성능 향상된 산출결과를 보았다. 이는 이동정보서비스의 현재 상황을 인식하기 위한 서비스 방법론에 매우 중요한 요소로 인식된다. 하지만 이러한 장점들에도 불구하고 정보검색 실패율이라는 취약점을 간과할 수는 없다.

본 논문에서 제시한 포아송 프리페칭은 어떤 특정 위치에서 사용자가 정보를 요구하였을 때, 최소의 MR(Missing Rate ; 정보검색 실패율)로 최대의 메모리를 활용할 수 있도록 하는 전략이다.

속도기반에서의 MR은 다른 알고리즘보다 가장 많은 프리페칭 영역을 설정하고 있으므로, 이에 대한 MT은 현저히 낮은 값을 가지므로, 우리는 이를 제외하고 빈도수기반의 프리페칭과 제안한 알고리즘의 MR을 비교한다.

< 표 2 > 정보검색 실패율

	빈도수기반의 프리패칭	포아송기반의 프리패칭
1	0.20	0.20
2	0.20	0.20
3	0.29	0.14
4	0.44	0.11
5	0.60	0.40
6	0.20	0.20
7	0.25	0.25
8	0.25	0.00
9	0.00	0.00
10	0.33	0.17
11	0.50	0.25
12	0.50	0.25
13	0.20	0.00
14	0.14	0.14
15	0.29	0.14
16	0.50	0.00
17	0.33	0.00
18	0.29	0.14
19	0.20	0.20
20	0.33	0.33



<그림 9> 정보검색 실패율

5. 결론

본 논문에서는 이동 환경에서 공간 지역성을 고려한 포아송분포를 적용시킨 프리패칭 방법을 제안하였다. 이는 기존의 속도 기반의 프리패칭과 빈도수 기반의 프리패칭 기법을 좀 더 향상시킨 방법을 제시하였다.

먼저 이동 시나리오를 제시하고 이에 따른 기존의 프리패칭 방법과 제안된 프리패칭 방법의 정보분할의 수와 정보의 실패율에 대해 비교 분석하였다. 이를 통해 제시한 프리패칭 방법이 이동 환경의 제약사항들을 효과적으로 극복할 수 있음을 알 수 있었다.

따라서 본 논문에서 제시한 프리패칭 전략은 메모리 활용을 극대화하면서 정보검색 실패율을 최소화시키는 좀 더 유연성 있는 전략임을 알 수 있었다.

참고문헌

1. 김문자, 차우석, 조인준, 조기환(2001), "위치인식 이동정보서비스를 위한 프리패칭 방법론," *정보처리학회 논문지*, 8-C , no.6, 12, 831-838.
2. B. Bruegge, B. Bennington. (1996), "Application of Mobile Computing and Communication," *IEEE Personal Communications*, 3 , no.1.
3. Dhananjay S. Phatak, and Rory Mulvaney(2002), "Clustering for Personalized Mobile Web Usage", *Proceedings of the IEEE FUZZ'02*.
4. Seungmin Park, Daeyoung Kim, Gihwan Cho(2004), "Improving prediction level of prefetching for location-aware mobile information service", *Future Generation Computer Systems archive* Vol 20 , 197 - 203.
5. Liangzhong Yin, Guohong Cao, Chita R. Das, Ajeesh Ashraf(2002)," Power-Aware Prefetch in Mobile Environments", *ICDCS* , 571-578
6. Tomasz Imielinski , Henny F. Korth(1996), "Introduction to Mobile Computing",*Mobile Computing, Kluwer-Academic Publishers*, 1-43.
7. Uwe Kubach, Kurt Rotherme(2001), "Exploiting location information for infostation-based hoarding", *Proceedings of the 7th annual international conference on Mobile computing and networking table of contents*, 15 - 27.
8. W. C. Peng and M. S. Chen(2000), "Mining User Moving Patterns for Personal Data Allocation in a Mobile Computing System ", *Proc. of the 29th International Conference on Parallel Processing* , 573-580.

[2006년 9월 접수, 2006년 11월 채택]