



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2007-0103800
(43) 공개일자 2007년10월25일

(51) Int. Cl.

G06F 15/16(2006.01) G06F 17/50(2006.01)

(21) 출원번호 10-2006-0035565

(22) 출원일자 2006년04월20일

심사청구일자 2006년04월20일

(71) 출원인

고려대학교 산학협력단

서울 성북구 안암동5가1 고려대학교 내

(72) 발명자

최봉대

서울 동대문구 이문2동 264-473

박철근

충남 천안시 쌍용동 월봉일성아파트 501동 1205호
(뒷면에 계속)

(74) 대리인

권혁성

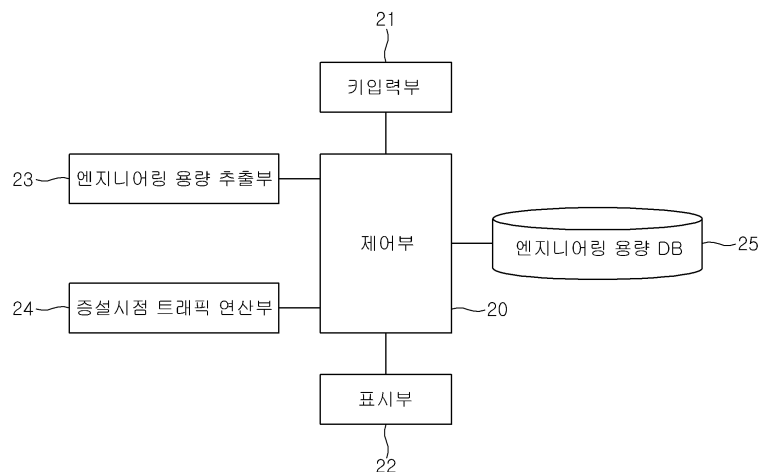
전체 청구항 수 : 총 13 항

(54) WAP G/W 서버 대수 예측 산정장치 및 그 동작방법

(57) 요약

본 발명은 증설시점에 필요한 WAP G/W 서버 대수를 예측 산정하는 장치 및 그 동작방법에 관한 것으로서, 증설시점의 트래픽 양을 예측하여 이를 이용하여 소정의 연산을 수행하여 얻은 WAP G/W 서버 대수를 디스플레이함을 특징으로 한다. 본 발명은, WAP G/W 서버 한 대에서 정상적으로 트래픽 처리할 수 있는 적정 CPU 점유율을 가질 때의 트랜잭션 용량인 엔지니어링 용량을 저장한 엔지니어링 용량 데이터베이스와, WAP G/W 서버에서 처리하는 트랜잭션의 오브젝트(object) 크기 분포 그래프를 이용하여 상기 엔지니어링 용량을 추출하는 엔지니어링 용량 추출부와, 소정의 공식에 의하여 WAP G/W 서버의 증설 시점의 총 트랜잭션 수를 연산하는 증설시점 트래픽 연산부와, 소정의 공식에 의하여 WAP G/W 서버의 증설 시점의 필요한 서버 대수를 예측 연산하는 제어부와, WAP G/W 서버 대수 예측 산정을 위하여 사용자로부터 관련 정보를 입력받는 키입력부와, 상기 연산된 증설 시점의 총 트랜잭션 수 및 증설시점의 필요한 서버대수를 디스플레이하는 표시부를 포함한다.

대표도 - 도2



(72) 발명자

황강욱

대전 서구 둔산동 둔지아파트 102-903호

이재문

서울 중구 을지로2가 11번지

윤종필

서울 중구 을지로2가 11번

특허청구의 범위

청구항 1

WAP G/W 서버 한 대에서 정상적으로 트래픽 처리할 수 있는 CPU 점유율을 가질 때의 트랜잭션 용량인 엔지니어링 용량을 저장한 엔지니어링 용량 데이터베이스와,

WAP G/W 서버에서 처리하는 트랜잭션의 오브젝트 크기 분포 그래프를 이용하여 상기 엔지니어링 용량을 추출하는 엔지니어링 용량 추출부와,

하기 청구항 식1에 의하여 WAP G/W 서버의 증설 시점의 총 트랜잭션 수를 연산하는 증설시점 트래픽 연산부와,

하기 청구항 식2에 의하여 WAP G/W 서버의 증설 시점의 필요한 서버 대수를 예측 연산하는 제어부

-청구항 식1-

증설 시점에서의 총 트랜잭션 수 (λ) = $A * U * Y$

(A: 총 가입자에 대한 유효가입자 비율, U: 증설시점의 총 가입자수, Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수)

-청구항 식2-

증설시점의 WAP G/W 서버 대수(N) = (λ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수

(λ : 증설시점에서의 총 트랜잭션 수)

를 구비한 WAP G/W 서버 대수 예측 산정장치.

청구항 2

WAP G/W 서버 한 대에서 정상적으로 트래픽 처리할 수 있는 적정 CPU 점유율을 가질 때의 트랜잭션 용량인 엔지니어링 용량을 저장한 엔지니어링 용량 데이터베이스와,

WAP G/W 서버에서 처리하는 트랜잭션의 오브젝트(object) 크기 분포 그래프를 이용하여 상기 엔지니어링 용량을 추출하는 엔지니어링 용량 추출부와,

하기 청구항 식1에 의하여 WAP G/W 서버의 증설 시점의 총 트랜잭션 수를 연산하는 증설시점 트래픽 연산부와,

하기 청구항 식2에 의하여 WAP G/W 서버의 증설 시점의 필요한 서버 대수를 예측 연산하는 제어부와,

WAP G/W 서버 대수 예측 산정을 위하여 사용자로부터 관련 정보를 입력받는 키입력부와,

상기 연산된 증설 시점의 총 트랜잭션 수 및 증설시점의 필요한 서버대수를 디스플레이하는 표시부

-청구항 식1-

증설 시점에서의 총 트랜잭션 수 (λ) = $A * U * Y$

(A: 총 가입자에 대한 유효가입자 비율, U: 증설시점의 총 가입자수, Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수)

-청구항 식2-

증설시점의 WAP G/W 서버 대수(N) = (λ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수

(λ : 증설시점에서의 총 트랜잭션 수)

를 구비한 WAP G/W 서버 대수 예측 산정장치.

청구항 3

제1항 또는 제2항에 있어서, 증설시점 트래픽 연산부는,

증설 시점에서의 총 트랜잭션 수 (λ) = $S * Y$

(Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S는 유효가입자수 일별 통계자료)

의 수학적식에 의하여 WAP G/W 서버의 증설 시점의 총 트랜잭션 수를 연산하는 WAP G/W 서버 대수 예측 산정장치.

청구항 4

제1항 또는 제2항에 있어서, 제어부는,

증설시점의 WAP G/W 서버 대수(N) = (S * Y / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수

(Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S: 유효가입자수 일별 통계자료)

의 수학적식에 의하여 WAP G/W 서버의 증설 시점의 필요한 서버 대수를 예측 연산하는 WAP G/W 서버 대수 예측 산정장치.

청구항 5

제1항 또는 제2항에 있어서, 상기 엔지니어링 용량 데이터베이스 내의 엔지니어링 용량은 관리자에 의해 수시로 업데이트 가능함을 특징으로 하는 WAP G/W 서버 대수 예측 산정장치.

청구항 6

제1항 또는 제2항에 있어서, 상기 적정 CPU 점유율은, 70% CPU 점유율을 가짐을 특징으로 하는 WAP G/W 서버 대수 예측 산정장치.

청구항 7

제1항 또는 제2항에 있어서, 상기 엔지니어링 용량 추출부에서 엔지니어링 용량을 추출하는 것은 오브젝트(object) 크기 분포 그래프에서 70% CPU 점유율을 가지도록 트랜잭션 수를 추출함을 특징으로 하는 WAP G/W 서버 대수 예측 산정장치.

청구항 8

WAP G/W 서버 한 대에서 정상적으로 트래픽 처리할 수 있는 적정 CPU 점유율을 가질 때의 트랜잭션 용량인 엔지니어링 용량을 추출하는 과정과,

WAP G/W 서버의 증설시점에 예측되는 트랜잭션을 소정의 트랜잭션 추출 연산식에 의하여 연산하는 과정과,

상기 엔지니어링 용량 및 상기 증설시점에 예측되는 트랜잭션을 이용하여 소정의 서버대수 연산식에 의하여 증설시점의 필요한 서버대수를 연산하는 과정

을 포함하는 WAP G/W 서버 대수 예측 산정 방법.

청구항 9

제8항에 있어서, 상기 연산되는 엔지니어링 용량 및 상기 증설시점에 예측되는 트랜잭션 및 증설시점의 필요한 서버대수를 디스플레이하는 과정을 더 포함하는 WAP G/W 서버 대수 예측 산정 방법.

청구항 10

제8항에 있어서, 상기 트랜잭션 추출 연산식은,

증설 시점에서의 총 트랜잭션 수 (λ) = A * U * Y ;

(A: 총 가입자에 대한 유효가입자 비율, U: 증설시점의 총 가입자수, Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수)

임을 특징으로 하는 WAP G/W 서버 대수 예측 산정 방법.

청구항 11

제8항에 있어서, 상기 트랜잭션 추출 연산식은,

증설 시점에서의 총 트랜잭션 수 (λ) = $S * Y$;

(Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S는 유효가입자수 일별 통계자료)

임을 특징으로 하는 WAP G/W 서버 대수 예측 산정 방법.

청구항 12

제8항에 있어서, 서버대수 연산식은,

증설시점의 WAP G/W 서버 대수(N) = (λ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수; (λ : 증설시점에서의 총 트랜잭션 수)

임을 특징으로 하는 WAP G/W 서버 대수 예측 산정 방법.

청구항 13

제8항에 있어서, 서버대수 연산식은,

증설시점의 WAP G/W 서버 대수(N) = ($S * Y$ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수 ;

(Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S: 유효가입자수 일별 통계자료)

임을 특징으로 하는 WAP G/W 서버 대수 예측 산정 방법.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

- <17> 본 발명은 증설시점에 필요한 WAP G/W 서버 대수를 예측 산정하는 장치 및 그 동작방법에 관한 것으로서, 증설시점의 트래픽 양을 예측하여 이를 이용하여 소정의 연산을 수행하여 얻은 WAP G/W 서버 대수를 디스플레이함을 특징으로 한다.
- <18> WAP(Wireless Application Protocol)이란 셀룰러폰이나 무선호출기 등의 무선장치들을 인터넷(전자우편, 웹, 뉴스그룹 및 IRC 등)에 접속할 수 있도록 그 방법을 표준화한 통신 프로토콜들의 규격을 말한다. 특히, WAP G/W(Wireless Application Protocol Gateway) 서버는 WAP 1.2/2.0 국제표준규격에 따라 제작되어 인터넷과 무선네트워크의 연결 Gateway로서 무선인터넷 서비스의 관문역할을 수행하는데, WAP Gateway 1.2버전은 프로토콜 변환기능을 제공하며, WAP Gateway 2.0은 xHTML과 Wireless Profiled HTTP를 지원한다.
- <19> 상기 WAP G/W 서버를 가지는 무선인터넷 시스템 도를 도 1에 도시하였다.
- <20> 상기 도 1을 참조하면, WAP G/W 서버(11)는 무선단말기와 웹서버 상호 간에 요청되는 각 오브젝트를 처리하기 위한 트랜잭션을 처리하는 기능을 수행한다. 이하에서는, 상기 WAP G/W 서버(11)는 WAP 게이트웨이(WAP gateway), WAP 서버, WAP DB 서버를 모두 포함하는 개념으로 정의하기로 한다.
- <21> 일반적으로 상기 WAP G/W 서버(11)의 설치 대수는, 순간 트래픽 증가 및 특정 시스템 장애 등의 운용마진을 감안하여 적정한 대수로 설계되어 구비되어야 한다. 상기 WAP G/W 서버(11)의 설치 대수 설계가 잘못 산정될 시에, 무선인터넷 사용량이 증가하여 트랜잭션이 증가하게 되면 WAP G/W 서버(11)가 처리할 수 있는 트랜잭션 용량을 초과하여 WAP G/W 서버가 정상적으로 동작할 수 없는 문제가 발생한다.
- <22> 따라서 WAP G/W 서버(11)의 설치 대수 설계는 시스템의 안정화 차원에서 중요한 과제가 된다. 특히, 무선인터넷 사용량 증가로 인해, 다가오는 미래의 WAP G/W 서버에서 처리해야 하는 트랜잭션 증가량을 예측하여 WAP G/W 서버의 증설 시점 및 증설 대수를 설계하는 것은 중요하다. 특히, WAP G/W 서버 증설 대수 구입비용을 미리 예산편성해 놓을 수 있도록, 미래의 WAP G/W 서버의 필요 대수를 정확히 예측 산정하는 것은 중요가 과제가 되어 가고 있는 실정이다.

발명이 이루고자 하는 기술적 과제

<23> 상기의 문제점을 해결하고자 본 발명은 안출된 것으로서, 미래의 트랜잭션을 예측하여 이에 필요한 WAP G/W 서버의 설치대수를 예측 산정하는 장치 및 방법을 제공함을 목적으로 한다. 또한, WAP G/W 서버의 설치대수를 예측 산정에 필요한 공식을 제시함을 목적으로 한다.

발명의 구성 및 작용

<24> 상기 목적을 이루기 위하여 본 발명은, WAP G/W 서버 한 대에서 정상적으로 트래픽 처리할 수 있는 적정 CPU 점유율을 가질 때의 트랜잭션 용량인 엔지니어링 용량을 저장한 엔지니어링 용량 데이터베이스와, WAP G/W 서버에서 처리하는 트랜잭션의 오브젝트(object) 크기 분포 그래프를 이용하여 상기 엔지니어링 용량을 추출하는 엔지니어링 용량 추출부와, 소정의 공식에 의하여 WAP G/W 서버의 증설 시점의 총 트랜잭션 수를 연산하는 증설시점 트래픽 연산부와, 소정의 공식에 의하여 WAP G/W 서버의 증설 시점의 필요한 서버 대수를 예측 연산하는 제어부와, WAP G/W 서버 대수 예측 산정을 위하여 사용자로부터 관련 정보를 입력받는 키입력부와, 상기 연산된 증설 시점의 총 트랜잭션 수 및 증설시점의 필요한 서버대수를 디스플레이하는 표시부를 포함한다.

<25> 또한, 본 발명은 WAP G/W 서버 한 대에서 정상적으로 트래픽 처리할 수 있는 적정 CPU 점유율을 가질 때의 트랜잭션 용량인 엔지니어링 용량을 추출하는 과정과, WAP G/W 서버의 증설시점에 예측되는 트랜잭션을 소정의 트랜잭션 추출 연산식에 의하여 연산하는 과정과, 상기 엔지니어링 용량 및 상기 증설시점에 예측되는 트랜잭션을 이용하여 소정의 서버대수 연산식에 의하여 증설시점의 필요한 서버대수를 연산하는 과정을 포함한다.

<26> 본 발명의 상세한 설명에 앞서 본 발명의 설명의 편의를 위하여 다음과 같이 용어를 정의하도록 한다.

<27> 가. 트랜잭션(transaction) - WAP G/W 기준으로 오브젝트를 요청받고 응답하는 일련의 절차를 의미한다. 상기 트랜잭션을 트래픽 기술 파라미터의 하나로서 이용한다.

<28> 나. 오브젝트(object) 크기 - 하나의 트랜잭션마다 WAP G/W가 응답하는 패킷들의 크기로서, 트랜잭션의 수와 함께 고려되어야 할 또 하나의 기술 파라미터이다.

<29> 다. 트래픽 - 트랜잭션이 오브젝트를 요청받고 응답하는 일련의 절차 및 그 크기에 따라 WAP G/W에 부과하는 부하로 정의한다.

<30> 라. 월별 최빈 트래픽 - 트래픽 발생 건을 10분 단위로 점핑하여 추출했을 시에, 한 달 중 가장 트래픽이 많았을 때의 발생 건을 말한다.

<31> 마. 공급기준 - 트래픽 증가 추이 및 환경변화에 따른 안정적인 용량 공급을 위해 1년 6개월 단위('Y+0.6')로 증설되는 공급기준을 말한다.

<32> 바. 시설용량 - 콘텐츠의 크기, 응답시간 등 적절한 서비스 품질을 만족하는 조건에서 WAP G/W 서버가 처리할 수 있는 최대 트래픽 처리 용량을 말한다.

<33> 사. 엔지니어링 용량 - 순간 트래픽 증가 및 특정 시스템 장애 등 운용마진을 감안하여 CPU 점유율 기준 30%의 여유를 두어 CPU 점유율 70%에 대응하는 용량을 말한다.

<34> 아. TPS(Transaction Per Second) - 용량단위로서 초당 트랜잭션 처리건수를 말한다.

<35> 이하, 본 발명의 바람직한 실시 예들의 상세한 설명이 첨부된 도면들을 참조하여 설명될 것이다. 하기에서 각 도면의 구성요소들에 참조부호를 부가함에 있어 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명을 설명함에 있어 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략할 것이다. 그리고 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의 내려진 용어들로서 이는 본 발명의 기술 분야에 속하는 자의 일반적 관례에 따라서 달라질 수 있으며, 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

<36> 도 2는 본 발명의 실시 예에 따른 WAP G/W 서버 대수 예측 산정장치의 내부 구성블록도이다.

<37> 상기 도 2를 참조하면, 키입력부(21)는 다수의 숫자키, 문자키, 특수키로 되어 있어 사용자와의 인터페이스를 담당하며, 표시부(22)는 LCD-TFT 등의 액정 디스플레이 패널로 되어 있어, 연산된 WAP G/W 서버 한 대당 엔지니어링 용량, 증설시점의 트래픽 크기, 증설시점의 WAP G/W 서버 대수 등을 디스플레이하여 사용자가 이를 알 수 있도록 한다.

<38> 엔지니어링 용량 데이터베이스(25)는 WAP G/W 서버 한 대가 CPU 점유율 70%가 될 때 처리할 수 있는 트랜잭션

수의 용량인 서버 한 대당 엔지니어링 용량을 저장하고 있는 데이터베이스이다. 순간 트래픽 증가 및 특정 시스템 장애 등의 운용마진을 감안하여 CPU 점유율 기준 30%의 여유를 두어 CPU 점유율 70%에 대응하도록 설계된다. 따라서 상기 WAP G/W 서버 한 대당 엔지니어링 용량이란, WAP G/W 서버 한 대가 정상적으로 트래픽 처리할 수 있도록 CPU 점유율 70%가 될 때 처리할 수 있는 트랜잭션 수의 용량을 말한다. 상기 용량의 단위로는 초당 트랜잭션 처리 건수를 의미하는 TPS(Transaction Per Second) 단위가 사용된다. 상기 엔지니어링 용량 데이터베이스(25)에 저장되는 엔지니어링 용량은 관리자에 의해 수시로 업데이트 저장될 수도 있지만, 하기 엔지니어링 용량 추출부(23)에서 추출한 엔지니어링 용량 결과값이 저장될 수 있다.

<39> 엔지니어링 용량 추출부(23)는 WAP G/W 서버 한 대당 엔지니어링 용량을 연산하여 상기 엔지니어링 용량 데이터베이스에 저장하는 기능을 수행한다. 상기 연산은 WAP G/W 서버에서 처리하는 트랜잭션의 오브젝트 크기 분포 그래프를 이용하여 추출할 수 있는데, 이러한 추출 과정은 하기 도 3의 S31 단계에서 상세히 설명하기로 한다.

<40> 증설시점 트래픽 연산부(24)는 증설시점('Y+0.5)의 총 트래픽을 연산하는 기능을 수행하는 기능부로서, 증설시점에서의 총 트랜잭션 수 (λ) = $A * U * Y = S * Y$ (상기에서, A: 총 가입자에 대한 유효가입자 비율, U: 증설시점의 총 가입자수, Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S는 유효가입자수 일별 통계자료)에 의해 연산을 수행한다. 상기 연산식은 도 3의 S32단계 및 하기 [식 1]에서 상세히 설명하기로 한다.

<41> 제어부(20)는 OS 운영체제 하에 상기 각 기능부를 제어하여 증설시점의 필요한 서버대수를 예측 산정하는 알고리즘을 구비한다. 즉, 상기 엔지니어링 용량 데이터베이스(25) 내의 엔지니어링 용량을 추출하고, 상기 증설시점 트래픽 연산부(24)에서 연산된 증설시점의 총 트랜잭션 수를 이용하여 증설시점의 { WAP G/W 서버 대수(N) = (λ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수 } 라는 식을 이용하여 증설시점의 서버대수를 예측하게 된다. 이러한 예측 알고리즘을 하기 도 3과 함께 상술한다.

<42> 도 3은 본 발명의 실시 예에 따른 WAP G/W 서버 대수를 예측 산정하는 과정을 도시한 플로차트이다.

<43> 먼저, WAP G/W 서버 한 대당 엔지니어링 용량을 추출하는 과정(S31)을 가진다. 상기 용량의 단위로는 초당 트랜잭션 처리 건수를 의미하는 TPS(Transaction Per Second) 단위가 사용된다. WAP G/W 서버 한 대당 엔지니어링 용량을 산출하는 것은 트래픽 예측에 앞서 WAP G/W 서버 한 대가 정상적으로 처리할 수 있도록 CPU 점유율 70%를 가지는 트래픽 처리 용량을 산출하기 위함이다.

<44> 상기 추출되는 WAP G/W 서버 한 대당 엔지니어링 용량의 예시를 [표 1]에 도시하였는데, CPU 점유율이 70% 되도록 처리될 수 있는 각 오브젝트 크기별 트랜잭션이 기재되어 있다. 결국, 이러한 각 오브젝트 크기별 트랜잭션이 엔지니어링 용량으로 사용되는 것이다. 하기 [표 1]을 보면, CPU 총 점유율은 70%에 근접한 69.87%를 가지며 트랜잭션의 총합은 194 [TPS]가 됨을 알 수 있다.

<45> [표 1]

<46>

[CPU 총 점유율 : 69.87 % 기준]	
오브젝트 크기[K]	트랜잭션[TPS]
3 [K]	162 [TPS]
5 [K]	21 [TPS]
10 [K]	9 [TPS]
30 [K]	2 [TPS]
합: 194 [TPS]	

<47> WAP G/W 서버 한 대당 엔지니어링 용량 추출은, 엔지니어링 용량 데이터베이스에 저장되어 있는 데이터에서 추출하여 이용할 수 있는데, 상기 엔지니어링 용량 데이터베이스에 저장되는 엔지니어링 용량은, 관리자에 의해 수시로 업데이트 수정/등록될 수 있다.

<48> 한편, 상기 [표 1]의 서버 한 대당 엔지니어링 용량은, 상기와 같이 엔지니어링 용량 데이터베이스를 통해 추출하지 않고, CPU 총 점유율이 70%에 근접하도록 각 오브젝트 크기별 트랜잭션을 사용자로부터 직접 키입력부를 통해 입력 받을 수 있는데, 발명의 이해를 돕기 위하여 사용자가 서버 한 대당 엔지니어링 용량을 산출하는 방법을 상술한다.

<49> 서버 한 대당 엔지니어링 용량을 산출하기 위해서는, 우선, WAP G/W 서버가 처리하는 트랜잭션의 오브젝트 크기 분포를 알아야 하는데, 도 4a에 도시한 바와 같이 WAP G/W 서버에서 처리하는 트랜잭션의 오브젝트 크기 분포

그래프를 이용한다. 도 4a의 오브젝트 크기 분포 그래프는 오브젝트 크기를 32등급으로 나누어 1K ~ 32K까지로 정한 후, 1K 미만은 1K에 합산하고 32K 이상은 32K에 합산한 다음 Y값들을 정규화하고, Y축을 로그(log) 스케일로 하여 생성되는 그래프이다. 그 후, 상기 정규화된 32등급(1~32K)의 오브젝트 크기 분포로부터 오브젝트 크기를 4등급으로 나누어 3K, 5K, 10K, 30K로 결정한 후, Y값들을 정규화하여 도 4b와 같이 3K->0.74551, 5K->0.11664, 10K->0.082601, 30K->0.055249라는 4등급의 오브젝트 크기 분포 그래프를 얻을 수 있다.

<50> 한편, 도 5의 시설용량 보정치 {3K:600TPS, 5K:500, 10K:100TPS}를 이용하여 오브젝트 크기당 트랜잭션의 수를 다음 [표 2]와 같이 구할 수 있다.

<51> [표 2]

<52>	3K: 445.31	5K: 58.32	10K: 25.61	30K: 5.52
------	------------	-----------	------------	-----------

<53> 다시 도 5를 이용하여 오브젝트 크기별 트랜잭션의 수에 대응하는 CPU 점유율을 구하면 하기 [표 3]과 같다.

<54> [표 3]

<55>	3K: 81.17	5K: 23.85	10K: 16.01	30K: 9.49	CPU 점유율 -> 130.52
------	-----------	-----------	------------	-----------	----------------------

<56> 상기 [표 3]과 같이 CPU 점유율이 100%를 넘는다는 것은 불합리하기 때문에, 순간 트래픽 증가 및 특정 시스템 장애 등의 운용마진을 감안하여 CPU 점유율이 70% 되도록 오브젝트 크기당 트랜잭션의 수를 일정비율로 감소하면서 CPU 점유율이 70%가 되도록 적절한 트랜잭션 수들의 4쌍을 구한다.

<57> 즉, 오브젝트 크기당 트랜잭션의 수를 일정비율로 감소하면서 CPU 점유율이 70%가 되도록 적절한 트랜잭션 수를 구하게 되면, 상기 [표 1]과 같이 CPU 총 점유율은 69.87이고 트랜잭션의 총합은 약 194 [TPS]를 가지게 된다. 결과적으로 상기 값이 WAP G/W 서버 1대당 엔지니어링 용량이 된다.

<58> 한편, 상기에 트랜잭션 수를 구할 때 CPU 점유율이 70%되도록 트랜잭션 수를 결정하는데, 이러한 70% 기준 근거 이유를 하기 [표 4], [표 5], [표 6]과 함께 간단히 설명한다.

<59> [표 4]

<60>	오브젝트 크기		3K	5K	10K	30K	합계
	CPU 점유율	30%	25.5	3.5	1.5	0.5	31 TPS
		40%	55.6	7.3	3.2	0.7	67 TPS
		50%	87.8	11.5	5.0	1.1	105 TPS
		60%	123.2	16.1	7.1	1.5	148 TPS
		70%	162.1	21.1	9.3	2.0	194 TPS
		80%	204.2	26.6	11.7	2.5	245 TPS
		90%	250.5	32.7	14.3	3.1	301 TPS

<61> 상기 [표 4]에서 TPS 합계는 도 4b에 도시한 오브젝트 크기분포{3K:0.74551, 5K:0.1164, 10K:0.082601, 30K:0.055249}를 가중치로 이용하여 얻어진 결과이다. 상기 [표 4]에서 CPU 점유율 70%일 때의 총 트래픽 처리량은 194 TPS가 됨을 알 수 있다.

<62> [표 5]

<63>	오브젝트 크기	3K	5K	10K	30K	합계
------	---------	----	----	-----	-----	----

CPU 점유율	30%	92.5	13.2	7.9	5.5	119 TPS
	40%	236.3	33.6	20.1	13.9	304 TPS
	50%	390.3	55.5	33.2	23.0	502 TPS
	60%	499.2	70.9	42.5	29.5	642 TPS
	70%	581.2	82.6	49.5	34.3	748 TPS
	80%	649.4	92.3	55.3	38.3	835 TPS
	90%	710.1	100.9	60.5	41.9	913 TPS

<64> 마찬가지로, 상기 [표 5]에서 TPS 합계는 마찬가지로 소정의 오브젝트 크기분포{3K:0.68642, 5K:0.20964, 10K:0.030504, 30K:0.073442}를 가중치로 이용하여 얻어진 결과이다. 상기 [표 5]에서 CPU 점유율 70%일 때의 총 트래픽 처리량은 748 TPS가 됨을 알 수 있다.

<65> 하기 [표 6]은 상기 [표 4] 및 [표 5]에서의 시험 결과와 실제의 실측자료를 도시한 테이블이다.

<66> [표 6]

CPU 점유율	30%	40%	50%	60%	70%	80%	90%
1차 시험결과 [표 4]	31	67	105	148	194	245	301
2차 시험결과 [표 5]	119	304	502	642	748	835	913
실측 자료	135	161	186	198	214	228	242

<68> 상기 [표 6]을 보면, WAP G/W 서버가 200 TPS 트래픽을 처리할 경우, '실측자료'에 대응되는 CPU 점유율은 60%를 약간 상회하고 '1차 시험결과'에 의한 대응되는 CPU 점유율은 70%를 다소 상회하는 정도이며, '2차 시험결과'에 의한 대응되는 CPU 점유율은 35% 정도가 된다. 이것으로 미루어 보아 '2차 시험결과'는 '실측자료'에 비하여 CPU 점유율을 현격하게 저평가하므로 엔지니어링 트래픽 기준으로 선택하기 불가능한 것으로 볼 수 있다. 한편, '1차 시험결과'는 '실측자료'에 비해 200 TPS 주변에서 CPU 점유율을 다소 높게 평가하지만 225 TPS 주변에서는 비스하게 평가한다. 그리고 이 주변에서 대응되는 CPU 점유율은 75% 정도가 된다. 그러나 250 TPS 주변에서는 '1차 시험결과'는 '실측자료'에 비해 CPU 점유율을 다소 낮게 평가하는 것을 알 수 있으며, 이 주변에서 '1차 시험결과'에 의한 대응되는 CPU 점유율은 80%를 다소 상회하고 '실측자료'에 의한 대응되는 CPU 점유율은 100%에 거의 달하고 있다. 이러한 결과로 미루어 엔지니어링 용량을 결정하기 위한 CPU 점유율은 75% 미만에서 적절한 70%로 결정하는 것이 바람직함을 알 수 있다.

<69> 한편, 상기와 같이 WAP G/W 서버 1대당 엔지니어링 용량 추출(S31)이 이루어진 후에는, 현시점 대비 증설시점(일명, 'Y+0.5' 시점)의 트래픽을 연산(S32)하는 과정을 가진다. 현시점 대비 18개월 후인 증설 시점('Y+0.5')의 총 트랜잭션(λ)은 하기 [식 1]에 의하여 구할 수 있다.

<70> [식 1]

<71> 'Y+0.5' 시점에서의 총 트랜잭션 수 (λ) = $A * U * Y$

<72> $= S * Y$

<73> (상기에서, A: 총 가입자에 대한 유효가입자 비율, U: 증설시점의 총 가입자수, Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S는 유효가입자수 일별 통계자료)

<74> 상기 증설시점의 총 가입자수(U)는 회귀분석을 통하여 증설시점에서 총 가입자수를 예측할 수 있는데, 상기 회귀분석에 대하여 설명하기로 한다.

<75> 2004년 1월부터 2005년 8월까지 PTAS 자료 및 원천자료로부터 도출한 WAP 버전별 월별 실측 TPS 변동 추이, 총 가입자(단말기 총 수)에 대한 가입자 단말기당 TPH 변동 추이, 그리고 WAP 서비스에 해당하는 총단말기 수의 변동추이에 관련된 테이블을 도 6에 도시하였는데, 상기 도 6의 테이블을 참조하여 WAP G/W 서버에 대한 단말기당 TPH 값 예측치 모델을 MINITAB이라는 통계 패키지를 이용한 회귀분석 그래프를 도시하면 도 7a, 도 7b와 같다.

- <76> 도 7a 및 도 7b에서 최번 10분으로 환산하기 위해서 1.4를 곱했다. 아래 적합되는 식에서 변수 (t)의 값은 X-좌표를 의미하고 2004년 1월일 때 X-좌표는 '1' 이며, Y-좌표는 '0.59'가 되고 월이 증가할 때마다 X-좌표는 1씩 증가하며 대응하는 Y-좌표는 WAP 2.0의 단말기당 TPH값을 나타낸다. 같은 방법으로 나타낸 WAP G/W 2.0의 단말기당 TPH에 해당하는 좌표 쌍의 집합은 $\{(1,0.59), (2,0.59), (3,0.57), \dots, (20,0.39)\}$ 이다.
- <77> 도 7a는 WAP G/W 서버의 단말기당 TPH에 해당하는 좌표 쌍의 집합(그림에서 검은 색과 부호 ◆ 모양을 나타냄)을 이용하여 선형으로 적합한 그림을 나타내었다. 상기 도 7a에서 부호 ■ 모양은 실측치에 대한 적합 직선의 TPH를 나타내고 부호 ◆ 모양은 예측치에 대한 TPH를 나타낸다. 적합하는 직선식은 하기 [식 2]와 같이 $y(t)=0.726674-0.0118737t$ 이다. 이렇게 하여 현시점(2005년 8월 기준)에서 Y+0.5에 해당하는 18개월 후(2007년 2월)의 예측 값은 $y(38)=0.726674-0.0118737 \times 38=0.2755$ (TPH) 이다.
- <78> 마찬가지로 방법으로 WAP G/W 서버의 TPH에 해당하는 좌표쌍의 집합을 이용하여 하기 [식 3]의 지수함수로 회귀 분석하여 적합하는 곡선 식은 다음과 같고 이에 대한 그림은 도 7b에 잘 나타나 있다.
- <79> [식 2] 선형적합 : $y(t)=0.726674-0.0118737$ (TPH)
- <80> [식 3] 지수함수 곡선적합 : $y(t)=0.713830(0.981988)^t$
- <81> 상기 식들을 이용하여 현시점(2005년 8월 기준)에서 18개월 후 증설시점에서 예측값들은 각각 [식 2]에 의해 선형: $y(38)=0.2755$, [식 3]에 의해 지수함수: $y(38)=0.3578$ 이 된다. 증설시점의 예측치를 비교해보면 지수함수 예측치는 선형 예측에 비해 과대평가 하며 선형 예측치는 지수함수에 비해 과소평가 한다는 점을 알 수 있는데 WAP 2.0의 단말기당 TPH는 계속 감소하는 추세에 있으므로 지수함수 적합 곡선을 예측 모델로 권장한다.
- <82> 도 7b에서 부호 ■ 모양은 실측치에 대한 적합곡선의 TPH를 나타내고 부호 ◆ 모양은 예측치에 대한 적합곡선의 TPH를 나타낸다. 이렇게 하여 현시점(2005년 8월 기준)에서 증설시점(2007년 2월)의 예측 값은 $y(38)=0.713830(0.981988)^{38}=0.3578$ 이 된다. 앞에서도 언급 했듯이 지수함수 적합 곡선은 선형 적합보다 과대 평가한다. 이렇게 하여 WAP G/W 2.0에 대한 TPS값 예측 모델로 지수함수 곡선 적합식 $y(t)=a(b)^t$ 을 권고한다. 여기서 상수 a 와 b 는 현시점(엔지니어링을 실시하는 현재 월)마다 새로이 갱신되어야 하며 이는 최소자승법 등의 방법이나 MINITAB등의 통계패키지를 이용하여 간단하게 결정할 수 있다. 앞에서도 언급했듯이 2004년 1월이 $t=1$ 에 해당한다는 점을 유의해야 한다. 특별히 현시점이 2005년 8월이면 $a=0.713803$ 이고 $b=0.981988$ 이며 현시점에서 $t=20$ 이며 증설시점에서 $t=38$ 이 된다.
- <83> 한편, 상기 [식 1]에서 증설시점에서의 유효 가입자당 예측 트랜잭션 수(Y)는 회귀분석을 통하여 구할 수 있는데 이에 대하여 설명한다.
- <84> 도 6은 2004년 1월부터 2005년 8월까지의 WAP G/W 서버의 월별 단말기수에 대한 통계를 테이블을 도시하였는데, 이를 근거로 한 회귀분석 그래프 예시를 도 8a 및 도 8b에 도시하였다.
- <85> 변수 (t)의 값은 X-좌표를 의미하고 2004년 1월일 때 X-좌표는 '1' 이며, Y-좌표는 '16,381'이 되고 월이 증가할 때마다 X-좌표는 1씩 증가하며 대응하는 Y-좌표는 WAP 2.0의 단말기수를 나타낸다. 같은 방법으로 나타낸 WAP G/W 서버의 단말기수에 해당하는 좌표 쌍의 집합은 $\{(1,16,381), (2,16,591), \dots, (20,18,400)\}$ 이다.
- <86> 도 8a는 WAP G/W 2.0의 단말기수에 해당하는 좌표 쌍의 집합(그림에서 ◆표 모양을 나타냄)을 이용하여 선형으로 적합한 그림을 나타낸다. 상기 도 8a에서 부호 ■ 모양은 실측치에 대한 적합직선의 단말기 수를 나타내고 부호 ◆ 모양은 예측치에 대한 적합직선의 단말기수를 나타낸다. 하기 [식 4]의 선형적합 식은 $y(t)=16557.1+97.8594t$ 이다. 이렇게 하여 현시점(2005년 8월 기준)에서 Y+0.5에 해당하는 18개월 후(2007년 2월)의 예측 값은 $y(38)=16557.1+97.8594 \times 38=20276$ (TPS)이다. 같은 방법으로 WAP G/W 2.0의 TPS에 해당하는 좌표쌍의 집합을 이용하여 하기 [식 5]의 지수함수, [식 6]의 로지스틱 S-성장 곡선으로 회귀 분석하여 적합하는 각각의 곡선식들은 다음과 같고 이들에 대한 그림은 각각 도 8b와 도 8c에 잘 나타나 있다.
- <87> [식 4] 선형적합 : $y(t)=16557.1+97.8594t$
- <88> [식 5] 지수함수 곡선적합 : $y(t)=16572.6(1.00561)^t$
- <89> [식 6] S-성장 곡선적합 : $y(t)=10^5/(7.03298-1.09143(1.02022)^t)$

- <90> 상기 식들을 이용하여 현시점(2005년 8월 기준)에서 18개월 후 증설시점에서 예측값들은 각각 [식 4]에 의한 선형의 $y(38)=20276$, [식 5]에 의한 지수함수의 $y(38)=20498$, [식 6]에 의한 S-성장의 $y(38)=21288$ 이 된다. 증설시점의 예측치를 비교해보면 S-곡선 예측치는 다른 것들에 비해 과대평가 하며 선형 예측치는 다른 것에 비해 과소평가 한다는 점을 알 수 있다. WAP G/W 서버의 단말기수는 계속 성장하는 추세에 있으므로 로지스틱 S-성장 곡선을 예측 모델로 권장한다.
- <91> 상기 도 8c에서 부호 ■ 모양은 실측치에 대한 적합곡선의 단말기 수를 나타내고 부호 ◆ 모양은 예측치에 대한 적합곡선의 단말기 수를 나타낸다. 이렇게 하여 현시점(2005년 8월)에서 증설시점(2007년 2월)의 예측 값은 $y(38)=10^5/(7.03298-1.09143(1.02022)^{38})=21288$ 이 된다.
- <92> 앞에서도 언급 했듯이 로지스틱 S-성장 곡선은 지수함수와 선형 적합보다 과대평가하나 지수함수 모델은 S-성장 모델보다 예측 구간이 현 시점에 비해서 멀어질수록 과대 평가하는 경향이 있다. 이렇게 하여 WAP G/W 서버에 대한 단말기수 예측 모델로 S-성장 곡선 적합식 $y(t)=10^5/(a+b(c)^t)$ 을 권고한다. 여기서 상수 a, b 와 c 는 현시점(엔지니어링을 실시하는 현재 월)마다 새로이 갱신되어야 하며 이는 최소자승법 등의 방법이나 MINITAB등의 통계패키지를 이용하여 간단하게 결정할 수 있다. 앞에서도 언급했듯이 2004년 1월이 $t=1$ 에 해당한다는 점을 유의해야 한다. 특별히 현시점이 2005년 8월이면 $a=7.04295$, $b=-1.09143$ 이고 $c=1.02022$ 이며 현시점에서 $t=20$ 이며 증설시점에서 $t=38$ 이 된다.
- <93> 한편, 상기 [식 1]에 의하여 증설시점('Y+0.5')에서의 총 트랜잭션 수 (??)를 연산(S32)한 후에는, 증설시점의 WAP G/W 서버 대수(N)를 예측 연산하는 과정(S33)을 가진다. 상기 연산은 하기 [식 7]에 의하여 이루어진다.
- <94> [식 7]
- <95> 증설시점의 WAP G/W 서버 대수(N) = (λ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수
- <96> 또는,
- <97> 증설시점의 WAP G/W 서버 대수(N) = ($S * Y$ / 서버 한 대당 엔지니어링 용량) - 현재 사용하는 서버 대수
- <98> (상기에서, λ : 증설시점에서의 총 트랜잭션 수, Y: 증설시점에서의 유효 가입자당 예측 트랜잭션 수, S: 유효 가입자수 일별 통계자료)
- <99> 상기에서 λ 는 증설시점에서의 총 트랜잭션 수는 상기 [식 1]에서 구한 값을 이용한 것이며, 상기 서버 한 대당 엔지니어링 용량은 S31 단계에서 추출한 값이다. 예컨대, 상기 서버 한 대당 엔지니어링 용량이 상기 [표 1]에서 구한 바와 같이 194 [TPS]인 경우 서버 한 대당 엔지니어링 용량 값으로 194가 대입된다.
- <100> 결국, 본 발명의 WAP G/W 서버 대수 예측 산정장치의 제어부는, 최종적으로 상기 [식 7]에 의해 증설시점('Y+0.5')의 WAP G/W 서버 대수(N)를 예측할 수 있게 된다. 예측된 서버 대수 값은 WAP G/W 서버 대수 예측 산정장치의 표시부를 통하여 디스플레이(S34)되어 사용자가 이를 이용할 수 있게 된다.
- <101> 상술한 본 발명의 설명에서는 구체적인 실시 예에 관해 설명하였으나, 여러 가지 변형이 본 발명의 범위에서 벗어나지 않고 실시 될 수 있다. 따라서 본 발명의 특허 범위는 상기 설명된 실시 예에 의하여 정할 것이 아니고 특허청구범위뿐 아니라 균등 범위에도 미침은 자명할 것이다.

발명의 효과

- <102> 상기에서 기술한 바와 같이 본 발명은, WAP G/W 서버 대수 예측 장치를 이용하여 미래의 증설시점의 WAP G/W 서버 대수를 예측하여 증설 시점의 총 트랜잭션 수 및 증설시점의 필요한 서버대수를 디스플레이함으로써, 장래에 필요한 WAP G/W 서버 대수를 손쉽게 예측할 수 있는 효과가 있다. 또한 상기 예측을 통하여 WAP G/W 서버 구입을 위한 예산 등을 미리 계획할 수 있는 효과가 있다.

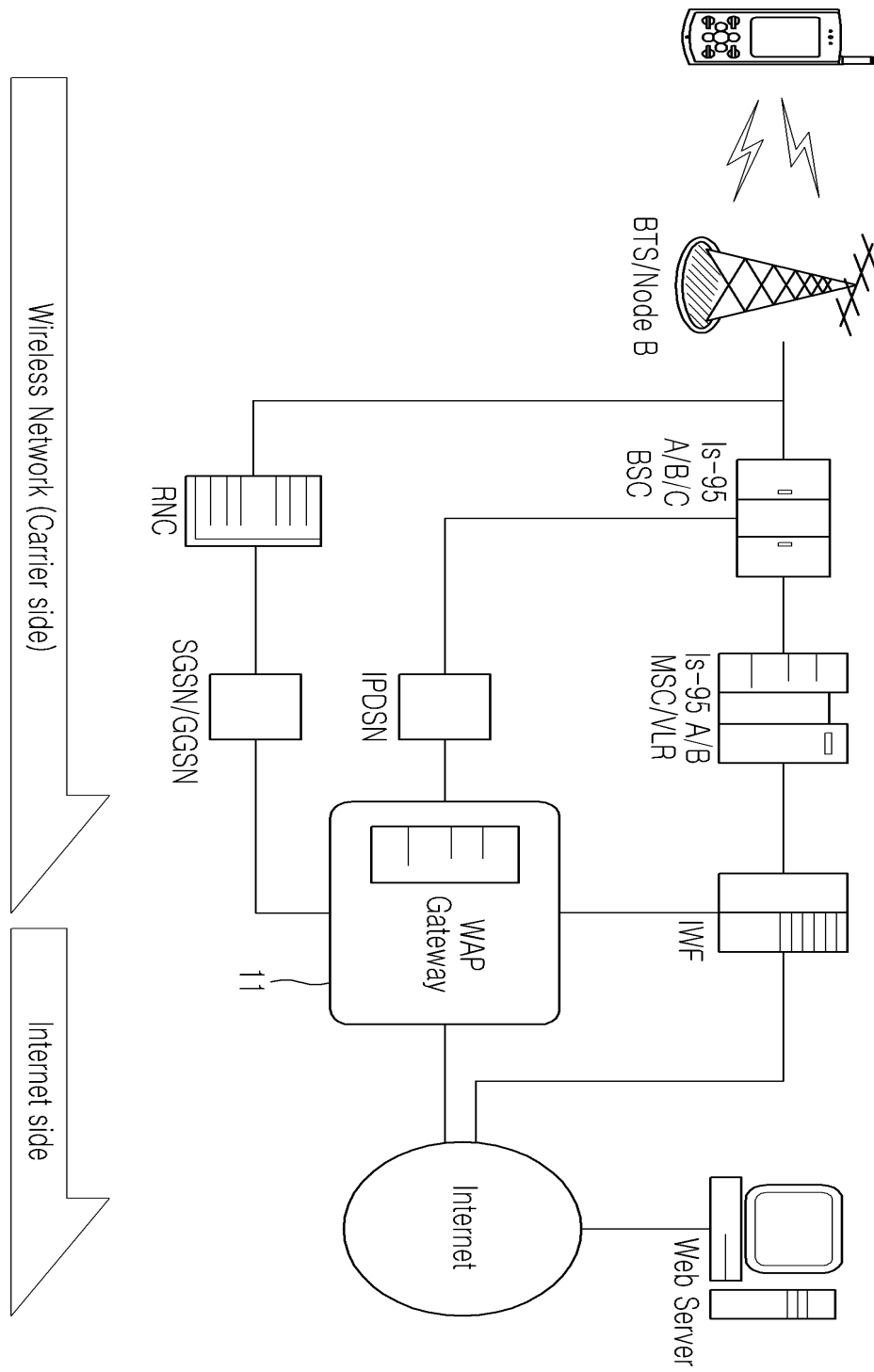
도면의 간단한 설명

- <1> 도 1은 일반적인 WAP G/W 시스템도이다.
- <2> 도 2는 본 발명에 따른 WAP G/W 서버 대수 예측 산정장치의 내부 구성 블록도이다.
- <3> 도 3은 본 발명에 따른 WAP G/W 서버 대수 예측 산정을 도시한 플로차트이다.

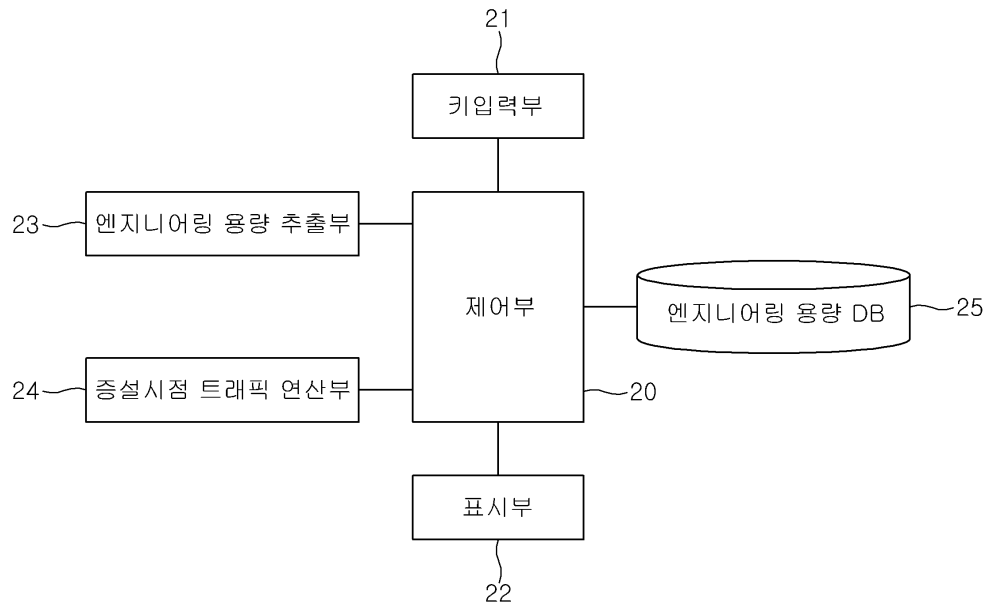
- <4> 도 4a는 정규화된 32등급(1~32K)의 WAP 오브젝트 크기 분포 그래프이다.
- <5> 도 4b는 4등급(3,5,10,30K)의 WAP 오브젝트 크기 분포 그래프이다.
- <6> 도 5는 각 오브젝트 크기별 CPU 점유율을 도시한 그래프이다.
- <7> 도 6은 가입자당 TPH 총가입자 월변동 추이를 나타낸 테이블도이다.
- <8> 도 7a는 월별 단말기당 TPH 추이를 선형함수 적합시켜 도시한 그래프이다.
- <9> 도 7b는 월별 단말기당 TPH 추이를 지수함수 적합시켜 도시한 그래프이다.
- <10> 도 8a는 월별 단말기수를 선형함수 적합시켜 도시한 그래프이다.
- <11> 도 8b는 월별 단말기수를 지수함수 적합시켜 도시한 그래프이다.
- <12> 도 8c는 월별 단말기수를 S-곡선 적합시켜 도시한 그래프이다.
- <13> *도면의 주요 부분에 대한 부호의 설명*
- <14> 20: 제어부 21: 키입력부
- <15> 22: 엔지니어링 용량 DB 23: 엔지니어링 용량 산출부
- <16> 24: 증설시점 트래픽 연산부 25: 엔지니어링 용량 DB

도면

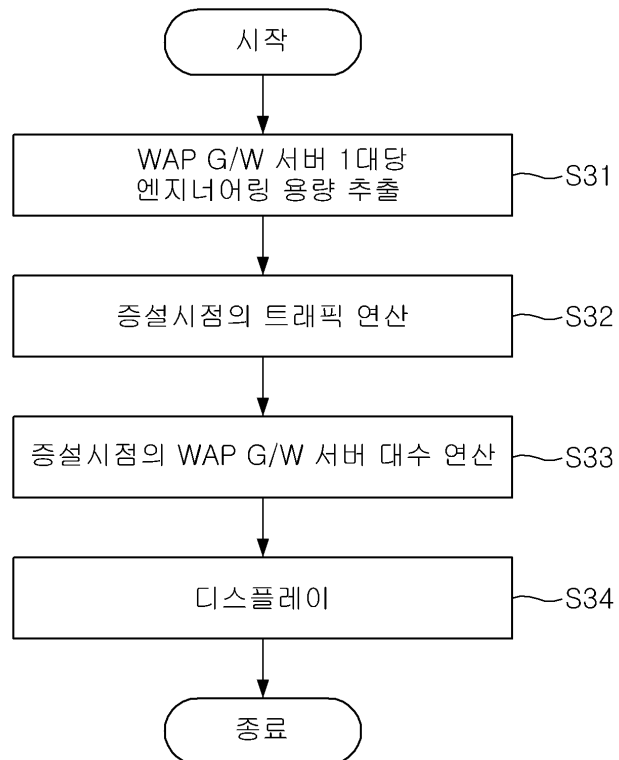
도면1



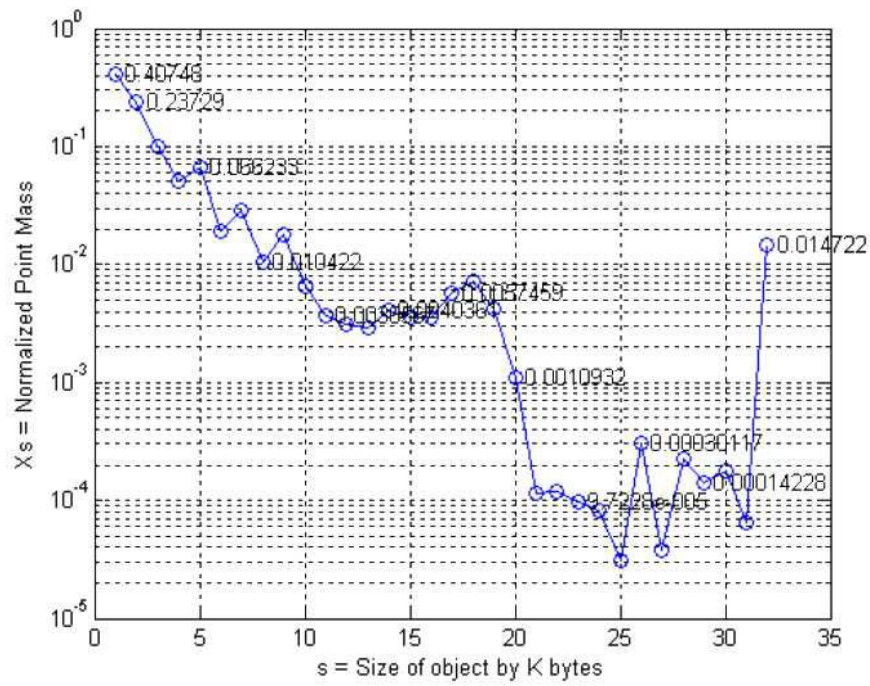
도면2



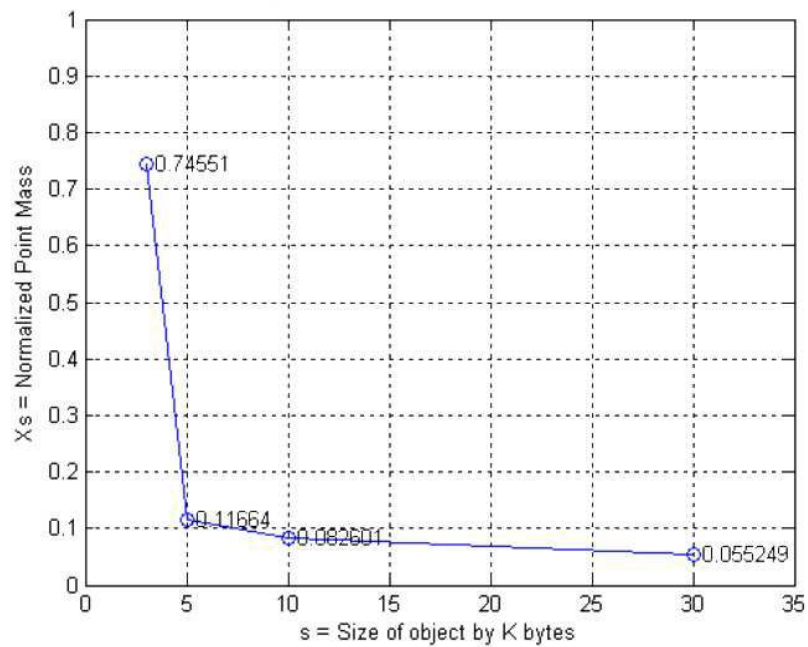
도면3



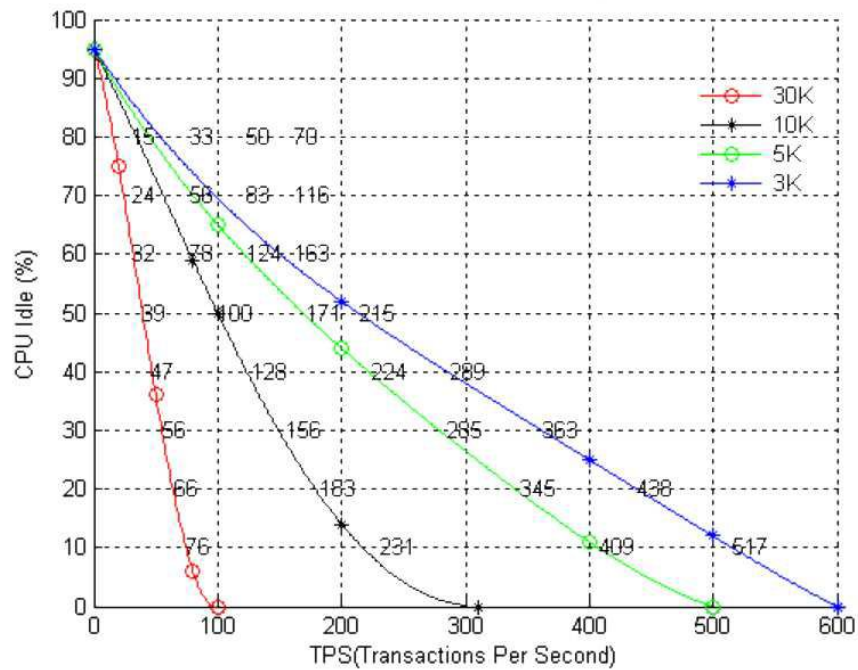
도면4a



도면4b



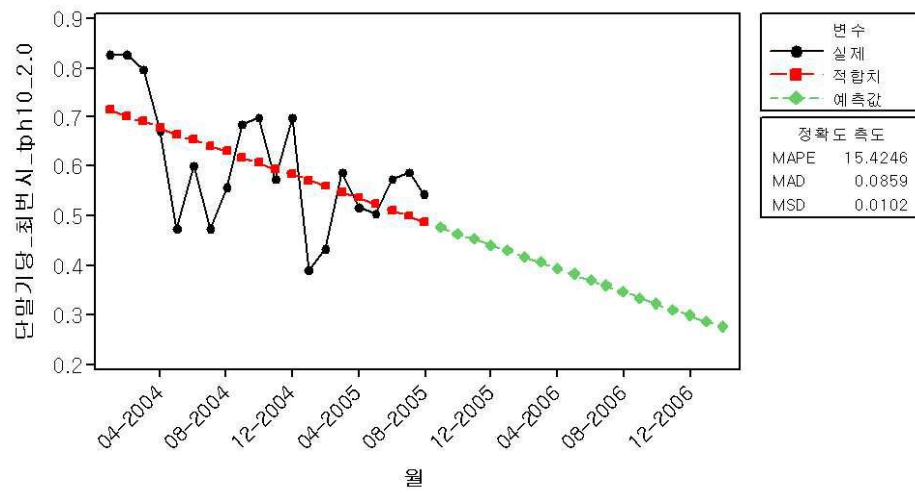
도면5



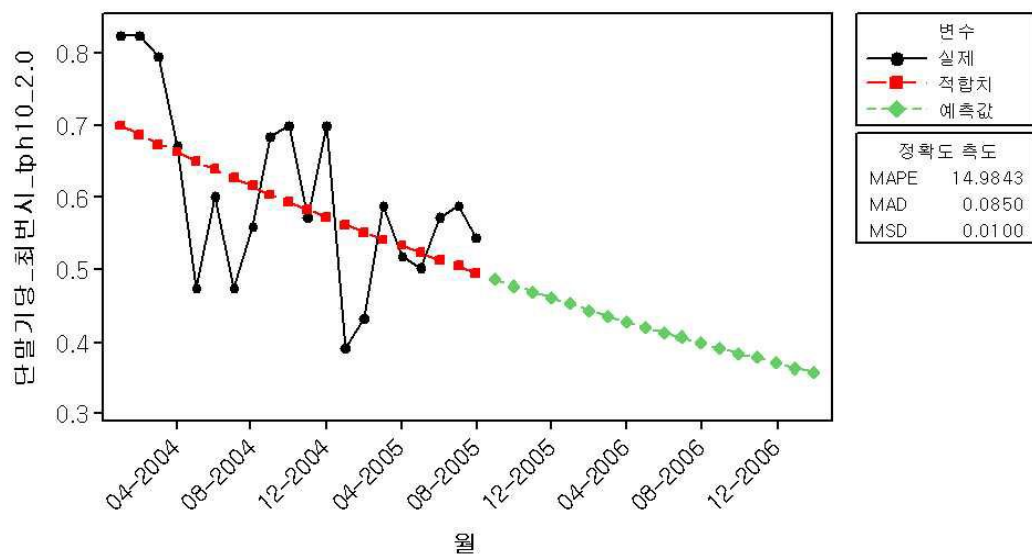
도면6

구분	실측 TPS	단말기당 TPH	단말기수(천)		비고
			A(22행)	B(21행)	
2004.1	250	0.59	16,381	17,074	
2004.2	300	0.59	16,591	17,222	
2004.3	351	0.57	16,784	17,355	
2004.4	340	0.48	16,977	17,493	
2004.5	281	0.34	17,176	17,645	
2004.6	396	0.43	17,228	17,667	
2004.7	362	0.34	17,500	17,901	
2004.8	467	0.40	17,564	17,937	
2004.9	582	0.49	17,452	17,799	
2004.10	633	0.50	17,517	17,843	
2004.11	542	0.41	17,612	17,920	
2004.12	692	0.50	17,669	17,958	
2005.1	422	0.28	17,846	18,114	
2005.2	503	0.31	17,956	18,207	
2005.3	741	0.42	18,059	18,293	
2005.4	683	0.37	18,141	18,375	
2005.5	721	0.36	18,214	18,451	
2005.6	864	0.41	18,283	18,526	
2005.7	942	0.42	18,343	18,596	
2005.8	921	0.39	18,400	18,665	

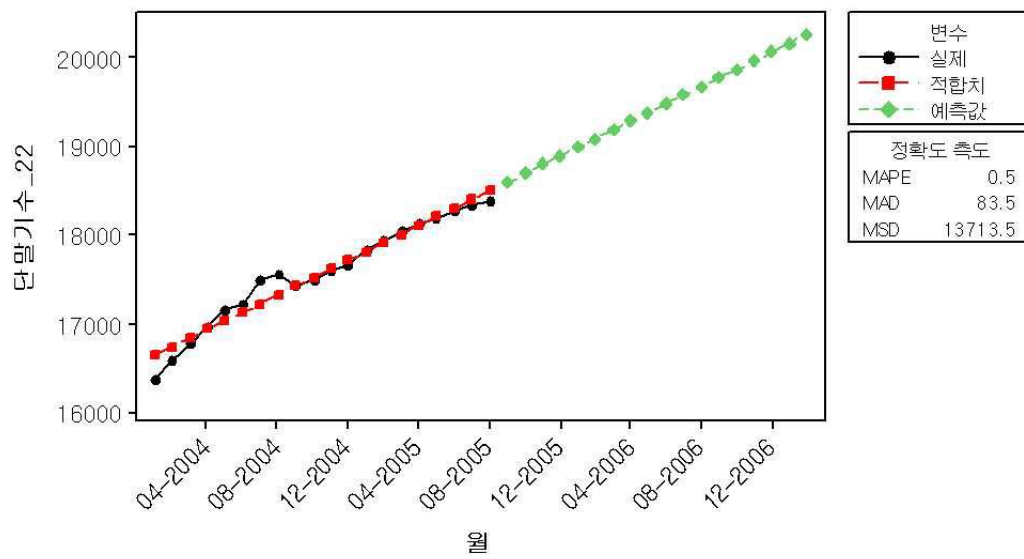
도면7a



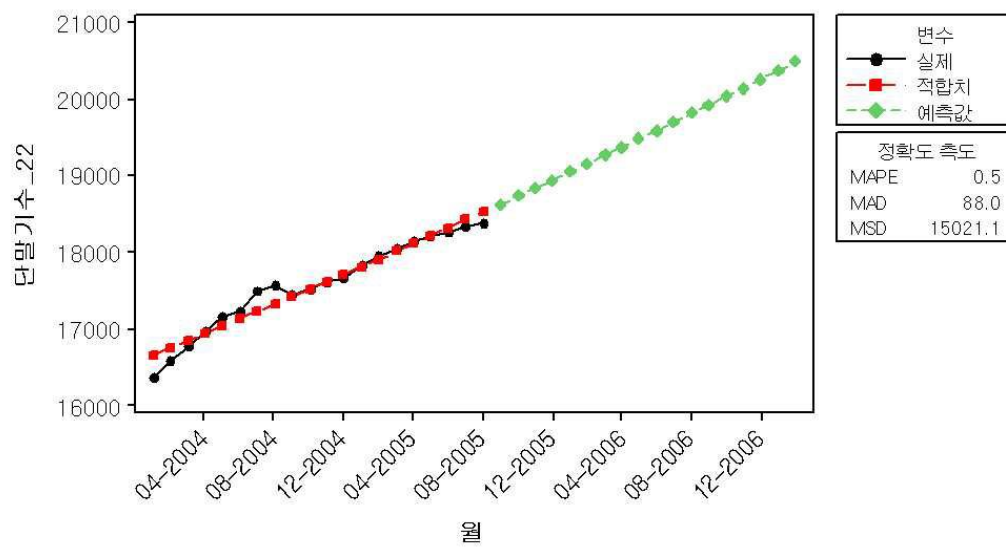
도면7b



도면8a



도면8b



도면8c

