

(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。 Int. Cl. G06F 17/00 (2006.01)	(45) 공고일자 (11) 등록번호 (24) 등록일자	2006년05월22일 10-0581084 2006년05월10일
---	-------------------------------------	--

(21) 출원번호 (22) 출원일자	10-2004-0104908 2004년12월13일	(65) 공개번호 (43) 공개일자
------------------------	--------------------------------	------------------------

(30) 우선권주장 1020040091585 2004년11월10일 대한민국(KR)

(73) 특허권자 한국전자통신연구원
 대전 유성구 가정동 161번지

 주식회사 케이티
 경기 성남시 분당구 정자동 206

(72) 발명자 최민석
 대전 서구 둔산동 아너스빌 918호

(74) 대리인 특허법인 신성

심사관 : 위재우

(54) 의사결정나무를 이용한 이메일 분류 장치 및 그 방법

요약

1. 청구범위에 기재된 발명이 속한 기술분야

본 발명은, 의사결정나무를 이용한 이메일 분류 장치 및 그 방법에 관한 것임.

2. 발명이 해결하려고 하는 기술적 과제

본 발명은, 클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류함으로써, 신속하고 간편하게 많은 수의 이메일을 분류하기 위한, 의사결정나무를 이용한 이메일 분류 장치 및 그 방법을 제공하는데 그 목적이 있음.

3. 발명의 해결방법의 요지

본 발명은, 의사결정나무를 이용한 이메일 분류 장치에 있어서, 폴더별로 이메일을 저장하기 위한 클라이언트 이메일 저장수단; 상기 클라이언트 메일 저장수단에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하기 위한 의사결정나무 생성수단; 이메일을 수신하기 위한 수신 이메일 처리수단; 상기 수신 이메일 처리수단에서 수신한 이메일을 저장하기 위한 이메일 저장수단; 상기 이메일 저장수단에 저장되어 있는 이메일을 의사결정나무를 이용하여 분류하기 위한 이메일 분류수단; 상기 이메일 분류수단에서 분류한 이메일을 상기 클라이언트 이메일 저장수단으로 전송하기 위한 이메일

일 전송수단; 및 상기 클라이언트 이메일 저장수단에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류하도록 상기 각 구성요소들을 제어하기 위한 제어수단을 포함함.

4. 발명의 중요한 용도

본 발명은 이메일 분류 장치 등에 이용됨.

대표도

도 5

색인어

의사결정나무, 폴더별 이메일 분류, 이메일 정보, 연관성

명세서

도면의 간단한 설명

도 1 은 본 발명에 따른 의사결정나무를 이용한 이메일 분류 장치의 일실시에 구성도,

도 2 는 본 발명에 따른 클라이언트 이메일 저장부의 폴더별 이메일 저장상태를 나타내는 일례시도,

도 3 은 본 발명에 따른 의사결정나무 생성 과정에 대한 일실시에 흐름도,

도 4 는 본 발명에 따른 의사결정나무의 일실시에 설명도,

도 5 는 본 발명에 따른 의사결정나무를 이용한 이메일 분류 방법에 대한 일실시에 흐름도이다.

* 도면의 주요 부분에 대한 부호의 설명

10 : 클라이언트 이메일 저장부 11 : 의사결정나무 생성부

12 : 제어부 13 : 기타 웹사이트

14 : 송신 이메일 처리부 15 : 수신 이메일 처리부

16 : 이메일 저장부 17 : 이메일 분류부

18 : 이메일 전송부

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 의사결정나무를 이용한 이메일 분류 장치 및 그 방법에 관한 것으로, 더욱 상세하게는 클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류하기 위한, 의사결정나무를 이용한 이메일 분류 장치 및 그 방법에 관한 것이다.

인터넷 및 웹의 발명과 함께 전세계인으로부터 널리 애용되고 있는 이메일은 정보통신 시대를 대표하는 응용프로그램(application program)이다. 21세기를 살아가는 현대인이라면 매일 소정의 이메일을 주고 받고 있으며, 자신이 가지고 있는 이메일 주소를 전화번호와 함께 중요한 연락용 매개체로 여기고 있다. 이렇듯 이메일에 의한 의사소통이 보편화됨에 따라 점차 이메일을 효과적으로 이용하고 관리하는 것이 중요한 과제로 인식되고 있다.

그런데 이메일을 이용하던 초기에는 웹이나 용량이 적은 클라이언트 프로그램을 이용해서 이메일을 송수신했기 때문에, 주고 받는 이메일을 체계적으로 분류하기보다는 일정시간마다 삭제하여 정리하는 것이 보다 중요한 과제였다. 설혹, 이메일을 분류해야 할 필요성이 생기더라도 이용자가 수동으로 정리하는 것만으로도 충분하였다.

그러나, 최근에는 개인이 대용량의 메일함을 구비하게 되고, 매일 수십 내지 수백통의 이메일을 수신함으로 인해서 메일함을 다양한 폴더로 구분하여 정리하는 이용자들이 증가하고 있다. 특히, 온라인 뉴스레터 등과 같은 푸쉬(push)형 서비스를 이용하고 있는 이용자에게는 폭증하는 이메일을 효율적으로 관리하는 것이 매우 중요한 과제로 인식되고 있다.

한편, 의사결정나무 학습 기법은 귀납추리(inductive inference)의 가장 대표적인 학습 기법으로 분류(classification)에 많이 쓰이는데, 일반적으로 이 학습 기법은 잡음(noise)에 강한 특징을 가지고 있다.

의사결정나무는 노드(node)로 구성되는데, 최상위에 위치하는 노드를 루트노드(root node)라 하고, 루트노드는 하위노드로 가지치기(pruning)하면서 노드의 수를 증가시킨다. 이 때, 루트노드와는 반대로 가장 하위에 위치하여 더 이상 가지치기가 진행되지 않는 노드를 종단노드(leaf node)라 한다. 그리고, 루트노드로부터 종단노드까지의 단계를 깊이(depth)라 한다.

이러한 의사결정나무 학습 기법은 수집된 데이터에 기초하여 트리 형태의 분류 모형을 형성하고, 이후 도착하는 데이터를 상기 분류 모형에 따라 분류하기 때문에 매우 우수한 자동 분류 방식으로 인식되고 있다.

한편, 종래의 규칙 기반의 이메일 자동 분류 방법은 가정-결론 방식으로, 이메일에 포함된 송신자 주소, 제목, 내용 등의 정보를 기초로 미리 정해진 규칙에 의해서 이메일을 해당 폴더로 분류한다. 즉, 송신자의 주소, 제목, 내용 등의 항목이 어떤 값을 가지면 어떤 폴더로 자동 분류하라는 규칙을 생성한 후 수신되는 메일들을 상기 설정한 규칙에 따라 분류한다.

예를 들어, "해가 뜨고 습도가 정상이라면 운동하러 간다" 라는 규칙과 "비가 오고 바람이 약하면 운동하러간다" 라는 규칙만을 가지고 있는 경우에, "해가 뜨고 습도가 정상인 날에는 자동적으로 운동하러 간다" 라는 결론에 도달할 수 있다. 반면에, 해가 뜨고 습도가 정상이 아닌 경우에는 해당하는 규칙이 없으므로 결론에 도달할 수 없다.

즉, 규칙 기반의 자동화는 이분법적인 결정이나 소수의 변수 및 경우의 수만이 존재할 때는 가장 간단하면서도 효과적인 자동화 방식이지만, 하기와 같은 문제점을 가지고 있다.

첫째, 처리해야 할 이메일이 증가하면 필요로 하는 규칙을 생성하고 관리하는데 소용되는 시간이 급격히 증가한다. 즉, 이메일을 분류하기 위해서는 각 항목들이 가지는 값의 수를 각각 곱한 만큼의 규칙이 필요한데, 변수와 경우의 수가 증가하면 굉장히 많은 규칙을 생성해야 하는 문제점이 있다.

둘째, 규칙 기반의 이메일 분류에서는 세밀한 분류를 위해서 설정하는 시간적 부담이 가중된다. 예를 들어, 하나의 이메일 주소로부터 발신된 메일을 각각 다른 폴더로 이동시키기 위해서는 메일 주소 외에 다른 변수를 고려해서 설정해 주어야 하는 문제점이 있다. 이 때, 설정한 규칙이 추가된 변수의 성격에 따라 유효하지 않게 작동할 수도 있다.

발명이 이루고자 하는 기술적 과제

본 발명은 상기 문제점을 해결하기 위하여 제안된 것으로, 클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류함으로써, 신속하고 간편하게 많은 수의 이메일을 분류하기 위한, 의사결정나무를 이용한 이메일 분류장치 및 그 방법을 제공하는데 그 목적이 있다.

본 발명의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있으며, 본 발명의 실시예에 의해 보다 분명하게 알게 될 것이다. 또한, 본 발명의 목적 및 장점들은 특허 청구 범위에 나타난 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

발명의 구성 및 작용

상기 목적을 달성하기 위한 본 발명의 장치는, 의사결정나무를 이용한 이메일 분류 장치에 있어서, 폴더별로 이메일을 저장하기 위한 클라이언트 이메일 저장수단; 상기 클라이언트 메일 저장수단에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하기 위한 의사결정나무 생성수단; 이메일을 수신하기 위한 수신 이메일 처리수단; 상기 수신 이메일 처리수단에서 수신한 이메일을 저장하기 위한 이메일 저장수단; 상기 이메일 저장수단에 저장되어 있는 이메일을 의사결정나무를 이용하여 분류하기 위한 이메일 분류수단; 상기 이메일 분류수단에서 분류한 이메일을 상기 클라이언트 이메일 저장수단으로 전송하기 위한 이메일 전송수단; 및 상기 클라이언트 이메일 저장수단에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류하도록 상기 각 구성요소들을 제어하기 위한 제어수단을 포함하는 것을 특징으로 한다.

한편, 본 발명의 방법은, 의사결정나무를 이용한 이메일 분류 방법에 있어서, 클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하는 의사결정나무 생성 단계; 외부로부터 이메일을 수신함에 따라 이메일 저장수단에 임시로 저장하는 임시 저장 단계; 상기 생성한 의사결정나무를 호출하여 연관 정도를 비교하는 비교 단계; 상기 비교 결과, 가장 높은 연관성을 가지는 폴더를 결정하는 폴더 결정 단계; 및 상기 결정한 폴더에 상기 수신한 이메일을 저장하는 이메일 저장 단계를 포함하는 것을 특징으로 한다.

상술한 목적, 특징 및 장점은 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 일실시예를 상세히 설명하기로 한다.

도 1 은 본 발명에 따른 의사결정나무를 이용한 이메일 분류 장치의 일실시예 구성도이다.

도 1 에 도시된 바와 같이, 본 발명에 따른 의사결정나무를 이용한 이메일 분류 장치는, 폴더별로 이메일을 저장하기 위한 클라이언트 이메일 저장부(10), 상기 클라이언트 메일 저장부(10)에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하기 위한 의사결정나무 생성부(11), 기타 웹사이트 등으로부터 이메일을 수신하기 위한 수신 이메일 처리부(15), 상기 수신 이메일 처리부(15)에서 수신한 이메일을 저장하기 위한 이메일 저장부(16), 상기 이메일 저장부(16)에 저장되어 있는 이메일을 의사결정나무를 이용하여 분류하기 위한 이메일 분류부(17), 상기 이메일 분류부(17)에서 분류한 이메일을 클라이언트 이메일 저장부(10)로 전송하기 위한 이메일 전송부(18), 클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류하도록 상기 각 구성요소들을 제어하기 위한 제어부(12)를 포함한다.

본 발명의 일실시예는 클라이언트로부터의 송신 이메일을 외부로 송신하기 위한 송신 이메일 처리부(14)를 더 포함한다.

여기서, 도 2 를 참조하여 클라이언트 이메일 저장부(10)의 폴더별 이메일 저장 방식에 대해 살펴보기로 한다.

도 2 에 도시된 바와 같이, 본 발명에 따른 클라이언트 이메일 저장부에는 수신 이메일을 분류/저장하기 위해 "CMS" 란 이메일 계정(ID)을 가진 클라이언트에 의해 다양한 폴더가 생성되어 있다.

즉, 최초 폴더는 "받은 편지함" 폴더와 "보낸 편지함" 폴더 등으로 구분하였고, 상기 "받은 편지함" 폴더는 "회사" 폴더, "광고" 폴더, "외부" 폴더, "중요 편지" 폴더 등으로 구분하였으며, 상기 "회사" 폴더는 "이전 업무" 폴더와 "2004년 업무" 폴더로 구분하였다.

따라서, 상기 "이전 업무" 폴더로 분류되는 메일은 "A001@major.etri.re.kr, 2003", "A002@major.etri.re.kr, 2002" 등을 포함할 수 있다. 즉, 상기 이메일은 받은 편지이고, 회사(etri)메일이며, 같은 부서(major)로부터 수신한 메일이며, 금년이 2004년이라 할 때 이전에 수신한 메일을 의미한다.

또한, 상기 "2004년 업무" 폴더로 분류되는 메일은 "A001@major.etri.re.kr, 2004", "A002@major.etri.re.kr, 2004" 등을 포함할 수 있다. 즉, 상기 이메일은 받은 편지이고, 회사(etri)메일이며, 같은 부서(major)로부터 수신한 메일이며, 2004년에 수신한 메일을 의미한다.

또한, 상기 "타 부서" 폴더로 분류되는 메일은 "A003@etri.re.kr", "A004@etri.re.kr" 등을 포함할 수 있다. 즉, 상기 이메일은 받은 편지이고, 회사(etri)메일이며, 타 부서로부터 수신한 메일을 의미한다.

또한, 상기 "외부" 폴더로 분류되는 메일은 "A005@kaist.ac.kr"을 포함할 수 있으며, 이는 받은 편지이고, 외부(kaist)로부터 수신한 메일을 의미한다.

또한, 상기 "중요 편지" 폴더로 분류되는 메일은 "A005@kaist.ac.kr, 문의"을 포함할 수 있으며, 이는 받은 편지이고, 중요(문의) 편지임을 의미한다.

이렇게 수신 이메일을 폴더별로 분류하고 해당 폴더에 이메일이 수신되었음을 표시함으로써, 클라이언트는 쉽게 해당 이메일이 수신되었음을 인지할 수 있다.

도 3 은 본 발명에 따른 의사결정나무 생성 과정에 대한 일실시에 흐름도이다.

먼저, 클라이언트 이메일 저장부(10)에는 클라이언트에 의해 폴더가 생성되어 있다(200).

이후, 상기 생성한 각 폴더에 해당 이메일을 이동시킨다(210).

이후, 상기 각 폴더와 해당 폴더에 저장되어 있는 메일을 기초로 의사결정나무를 생성한다(203).

이후, 상기 생성한 의사결정나무를 제어부(12)로 전달하여 저장하도록 한다(204).

이러한 과정을 통해 의사결정나무를 이용한 이메일 분류 준비를 완료한다.

도 4 는 본 발명에 따른 의사결정나무의 일실시에 설명도이다.

도 4 에 도시된 바와 같이, 본 발명에 따른 의사결정나무는, "메일서버 C 클래스"를 시작점으로 회사(일예로, etri.re.kr) 메일 여부에 따라 "메일서버 D 클래스"와 "제목"의 분류 패턴을 갖고, 상기 "메일 서버 D 클래스"는 부서(일예로 major.etri.re.kr)의 동일 여부에 따라 "기간"과 "다른 부서"의 분류 패턴을 갖는다.

또한, 상기 "기간"은 수신년도에 따라 이전업무와 2004년 업무의 분류 패턴을 갖는다.

또한, 상기 "제목"은 중요도(일예로 문의)에 따라 중요편지와 비중요편지의 분류 패턴을 갖는다.

본 발명의 일실시에에서는 3단계의 깊이를 갖는 의사결정나무를 예로 들어 설명하였지만, 깊이는 얼마든지 증가시킬 수 있으므로 깊이는 본 발명에 아무런 영향을 끼치지 않는다.

도 5 는 본 발명에 따른 의사결정나무를 이용한 이메일 분류 방법에 대한 일실시에 흐름도이다.

먼저, 도 3 에 도시된 바와 같은 과정을 통해 의사결정나무가 생성되어 있다.

이후, 외부로부터 이메일을 수신함에 따라 이메일 저장부(16)에 임시로 저장한다(301, 302).

이후, 기 생성한 의사결정나무를 호출하여 연관 정도를 비교한다(303, 304). 즉, 상기 수신한 이메일 정보와 의사결정나무를 비교한다.

이후, 상기 수신한 이메일의 저장 폴더를 결정한 후 클라이언트 이메일 저장부(10)로 전송하여 해당 폴더에 저장하도록 한다(305, 306). 이 때, 저장 폴더는 상기 비교결과 가장 높은 연관성을 가지는 폴더이다.

이후, 상기 이메일 저장부(16)에서 상기 임시 저장된 이메일을 삭제한다(307).

이상에서 설명한 본 발명은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니다.

발명의 효과

상기와 같은 본 발명은, 클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류함으로써, 신속하고 간편하게 많은 수의 이메일을 분류할 수 있는 효과가 있다.

또한, 본 발명은 클라이언트의 이메일 분류 패턴을 분석함으로써, 클라이언트에게 차별적인 부가 서비스를 제공할 수 있는 효과가 있다.

또한, 본 발명은 의사결정나무를 이용하여 이메일을 폴더별로 자동으로 분류하기 때문에 별도의 사용법을 학습을 받지 않아도 되므로 클라이언트의 편의성을 증대시킬 수 있는 효과가 있다.

(57) 청구의 범위

청구항 1.

의사결정나무를 이용한 이메일 분류 장치에 있어서,

폴더별로 이메일을 저장하기 위한 클라이언트 이메일 저장수단;

상기 클라이언트 메일 저장수단에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하기 위한 의사결정나무 생성수단;

이메일을 수신하기 위한 수신 이메일 처리수단;

상기 수신 이메일 처리수단에서 수신한 이메일을 저장하기 위한 이메일 저장수단;

상기 이메일 저장수단에 저장되어 있는 이메일을 의사결정나무를 이용하여 분류하기 위한 이메일 분류수단;

상기 이메일 분류수단에서 분류한 이메일을 상기 클라이언트 이메일 저장수단으로 전송하기 위한 이메일 전송수단; 및

상기 클라이언트 이메일 저장수단에 폴더별로 저장되어 있는 이메일 정보를 기초로 의사결정나무를 생성하고, 외부로부터 이메일을 수신함에 따라 상기 생성한 의사결정나무를 이용하여 분류하도록 상기 각 구성요소들을 제어하기 위한 제어수단

을 포함하는 의사결정나무를 이용한 이메일 분류 장치.

청구항 2.

제 1 항에 있어서,

상기 이메일 정보는,

이메일 주소, 수신날짜, 메일서버의 클래스 넘버, 메일 제목을 포함하는 것을 특징으로 하는 의사결정나무를 이용한 이메일 분류 장치.

청구항 3.

의사결정나무를 이용한 이메일 분류 방법에 있어서,

클라이언트가 생성한 폴더와 상기 폴더내에 저장되어 있는 이메일이 가지고 있는 정보를 기초로 의사결정나무를 생성하는 의사결정나무 생성 단계;

외부로부터 이메일을 수신함에 따라 이메일 저장수단에 임시로 저장하는 임시 저장 단계;

상기 생성한 의사결정나무를 호출하여 연관 정도를 비교하는 비교 단계;

상기 비교 결과, 가장 높은 연관성을 가지는 폴더를 결정하는 폴더 결정 단계; 및

상기 결정한 폴더에 상기 수신한 이메일을 저장하는 이메일 저장 단계

를 포함하는 의사결정나무를 이용한 이메일 분류 방법.

청구항 4.

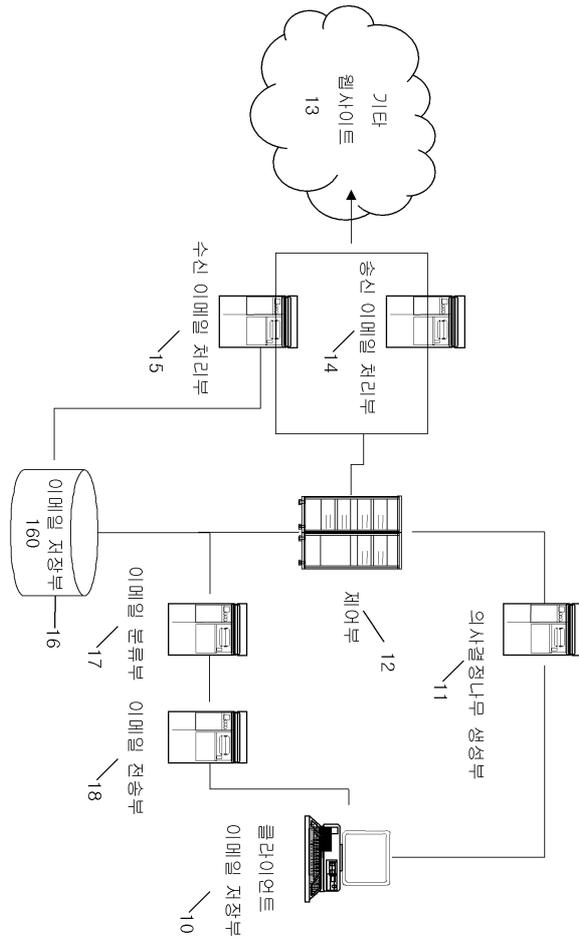
제 3 항에 있어서,

상기 이메일 정보는,

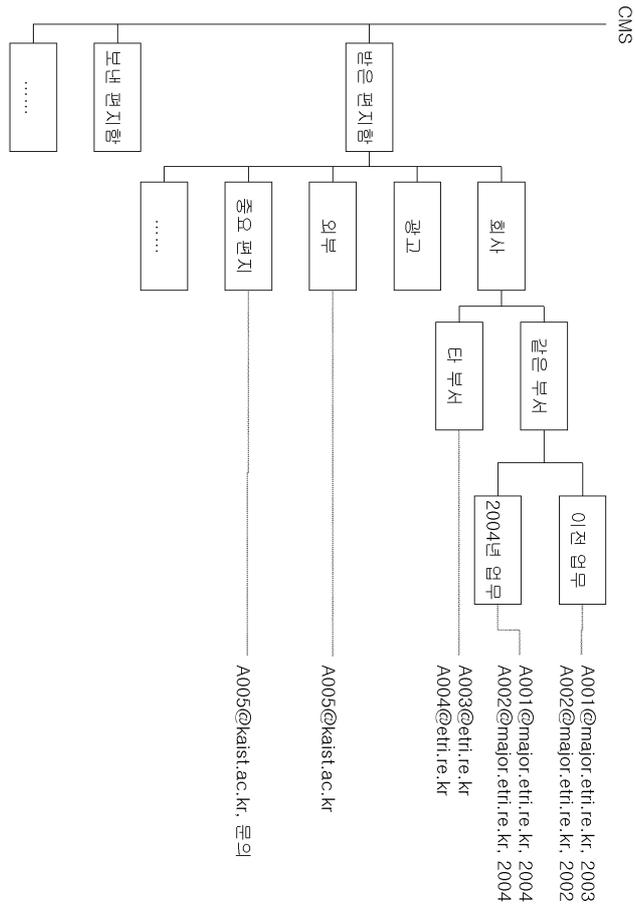
이메일 주소, 수신날짜, 메일서버의 클래스 넘버, 메일 제목을 포함하는 것을 특징으로 하는 의사결정나무를 이용한 이메일 분류 방법.

도면

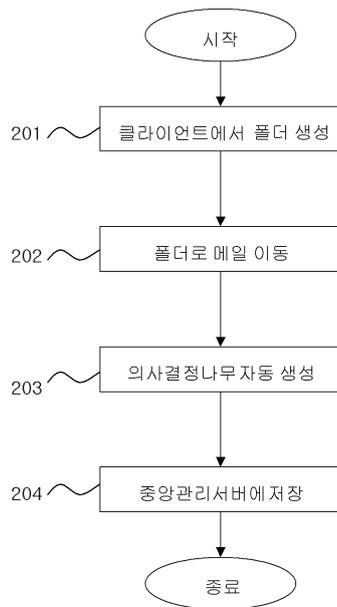
도면1



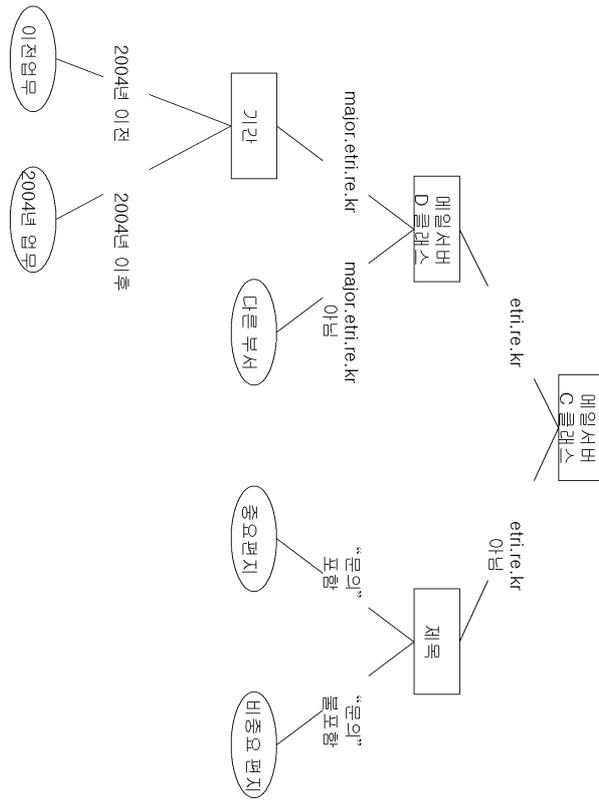
도면2



도면3



도면4



도면5

