(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2010-0013989
(43) 공개일자 2010년02월10일

(51) Int. Cl.
H04M 3/436 (2006.01) *H04M 3/42* (2009.01)
(21) 출원번호 10-2008-0075755
(22) 출원일자 2008년08월01일
심사청구일자 2008년08월01일

(71) 출원인
한국정보보호진흥원
서울특별시 송파구 가락동 78번지 IT벤처타워 서관
(72) 발명자
정종일
경기도 오산시 수청동 우미이노스빌아파트 109-1002
이태진
서울특별시 광진구 자양동 779-11 현대하이엘 1603호
(뒷면에 계속)
(74) 대리인
김영철, 김 순 영

전체 청구항 수 : 총 13 항

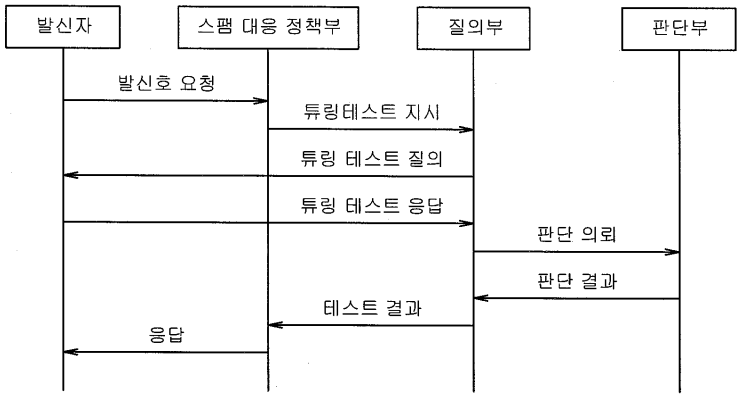
(54) VoIP 환경에서의 튜링 테스트 기반 스팸 차단 장치 및그 방법

(57) 요 약

발신호 요청을 하는 발신자에게 튜링 테스트의 질의 메시지를 전송하는 질의부; 발신자의 응답 메시지와 질의 메시지의 정답이 일치하는 지 비교하는 판단부; 및 판단부의 비교 결과를 근거로 발신자의 상태를 전이시키는 스팸 대응 정책부를 포함하는 스팸 차단 장치가 개시된다.

본 발명에 따르면, 종래의 기술에서 스패머로 분류하는 판단의 오탐율을 낮추고, 발신 전용의 소프트웨어를 이용한 대량의 스팸 전송을 차단할 수 있다. 또한 튜링 테스트의 결과를 발신자에게 전송하여, 정상적인 발신자에게 자신이 스패머로 분류 되었는지를 인지하게 할 수 있다.

대 표 도 - 도3



(72) 발명자

윤석웅

경기도 성남시 중원구 성남동 2396번지 신동아파밀리에 1003호

김중만

경기도 남양주시 화도읍 창현리 신명아파트 103동 1108호

원용근

서울특별시 송파구 가락동 4-5번지 e-house 403호

정현철

서울특별시 송파구 오금동 55-14 멀티파크 B동 201호

원유재

경기도 용인시 수지구 신봉동 LG빌리지5차아파트 515동 1401호

특허청구의 범위

청구항 1

스팸 차단 장치에 있어서,

발신호 요청을 하는 발신자에게 튜링 테스트의 질의 메시지를 전송하는 질의부;

상기 발신자의 응답 메시지와 상기 질의 메시지의 정답이 일치하는 지 비교하는 판단부; 및

상기 판단부의 비교 결과를 근거로 발신자의 상태를 전이시키는 스팸 대응 정책부를 포함하는 것을 특징으로 하는 스팸 차단 장치.

청구항 2

제1항에 있어서,

상기 질의부에서 전송하는 질의 메시지는 음성 또는 CAPTCHA에 의하여 제공되는 것을 특징으로 하는 스팸 차단 장치.

청구항 3

제1항에 있어서,

상기 스팸 대응 정책부는,

발신자의 상태가 스팸 발신자 후보로 분류된 상태(예비 블랙 상태)인 경우에 상기 질의부로 튜링 테스트 수행을 지시하는 것을 특징으로 하는 스팸 차단 장치.

청구항 4

제3항에 있어서,

상기 스팸 대응 정책부는,

상기 판단부의 비교 결과가 일치하는 경우에, 상기 발신자의 상태를 정상적인 발신자로 분류된 상태(그레이 상태)로 전이시키는 것을 특징으로 하는 스팸 차단 장치.

청구항 5

제3항에 있어서,

상기 스팸 대응 정책부는,

상기 판단부의 비교 결과가 일치하지 않는 경우에, 상기 발신자의 상태를 스팸 발신자로 분류된 상태(블랙 상태)로 전이시키는 것을 특징으로 하는 스팸 차단 장치.

청구항 6

제1항에 있어서,

상기 판단부의 비교 결과를 근거로 상기 발신자에게 발신호 요청에 대한 허용 여부를 알리는 메시지를 전송하는 인증부를 더 포함하는 것을 특징으로 하는 스팸 차단 장치.

청구항 7

제6항에 있어서,

상기 인증부는,

상기 판단부의 비교 결과가 일치하는 경우에는 발신자에게 발신호 요청에 대한 허용 메시지를 전송하고,

상기 판단부의 비교 결과가 일치하지 않는 경우에는 발신자에게 발신호 요청에 대한 불허 메시지를 전송하는 것을 특징으로 하는 스팸 차단 장치.

청구항 8

스팸 차단 방법에 있어서,

발신자의 상태를 판단하는 단계;

발신자의 상태가 스팸 발신자 후보로 분류된 상태(예비 블랙 상태)인 경우에 튜링 테스트 수행을 지시하는 단계;

발신자에게 튜링 테스트의 질의 메시지를 전송하는 단계;

상기 발신자의 응답 메시지와 상기 질의 메시지의 정답이 일치하는지 비교하는 단계; 및

상기 비교하는 단계의 결과를 근거로 발신자의 상태를 전이시키는 단계를 포함하는 것을 특징으로 하는 스팸 차단 방법.

청구항 9

제8항에 있어서,

상기 질의 메시지는 음성 또는 CAPTCHA에 의하여 제공되는 것을 특징으로 하는 스팸 차단 방법.

청구항 10

제8항에 있어서,

상기 발신자의 상태를 전이시키는 단계는,

상기 비교하는 단계의 결과가 일치하는 경우 상기 발신자의 상태를 정상적인 발신자로 분류된 상태(그레이 상태)로 전이시키는 것을 특징으로 하는 스팸 차단 방법.

청구항 11

제8항에 있어서,

상기 발신자의 상태를 전이시키는 단계는,

상기 비교하는 단계의 결과가 일치하지 않는 경우 상기 발신자의 상태를 스팸 발신자로 분류된 상태(블랙 상태)로 전이시키는 것을 특징으로 하는 스팸 차단 방법.

청구항 12

제8항에 있어서,

상기 비교하는 단계의 결과를 근거로 상기 발신자에게 발신호 요청에 대한 허용 여부를 알리는 메시지를 전송하는 단계를 더 포함하는 것을 특징으로 하는 스팸 차단 방법.

청구항 13

제12항에 있어서,

상기 메시지를 전송하는 단계는,

상기 비교하는 단계의 결과가 일치하는 경우에는 발신자에게 발신호 요청에 대한 허용 메시지를 전송하고,

상기 비교하는 단계의 결과가 일치하지 않은 경우에는 발신자에게 발신호 요청에 대한 불허 메시지를 전송하는 단계를 포함하는 것을 특징으로 하는 스팸 차단 방법.

명 세 서

발명의 상세한 설명

기 술 분 야

[0001] 본 발명은 튜링 테스트 기반 스팸 차단 장치 및 그 방법에서, 그레이리스트 기반 필터링을 기초로 하여, 특정

상태에 있는 발신자를 대상으로 튜링 테스트를 시행하고, 튜링 테스트의 결과를 반영하여 발신자의 상태를 전이시키는 튜링 테스트 기반 스팸 차단 장치 및 그 방법에 관한 것이다.

배 경 기 술

[0002] VoIP 스팸은 VoIP 단말에서 발신 및 수신되거나, VoIP 통신채널을 통해 전달되는 불법적인 음성이나 문자 스팸을 의미한다.

[0003] VoIP 스팸은 스팸 컨텐츠의 내용에 따라서 크게 전화스팸(SPIT, SPam over Internet Telephony)와 문자스팸(SPIM, SPam over Instant Messaging)으로 분류된다.

[0004] 여기서, 전화스팸은 수신자와 연결을 맺고 나서 스팸 호를 전송하는 방식이며, 문자스팸은 호 연결을 위한 시그널 메시지에 스팸메시지를 담아 수신자가 호에 응답하기 전에 스팸메시지를 텍스트 형태로 보여주는 방식이다. 문자스팸의 경우, 스패머(spamer)는 수신자가 호 수락 전에 통화를 중단시켜 요금청구를 회피할 수 있다.

[0005] 추가적으로, VoIP 스팸은 스팸 전달 방식에 따라, 정상적인 통화절차로 발신하는 방법과 수신단말로 직접 발신하는 방법으로 구분할 수 있다. 전자는 일반적으로 발신자와 수신자 간의 정상적인 호 전달 과정과 동일하다. 후자는 발신자가 수신자의 주소 및 전화번호를 알 수 있다면, 수신단말로 바로 호를 연결하는 P2P 방식 등을 사용하여 스팸을 전송할 수 있다.

[0006] 종래의 스팸대응 기술에 있어서, 이메일스팸에 대한 스팸대응 기술은 이메일스팸이 텍스트 기반이므로 컨텐츠 필터링 등을 통해 스팸 수신 전에 차단이 가능하다.

[0007] 그러나, VoIP 스팸의 경우는, 전화서비스 특성상 불특정 다수로부터 수신되는 호를 사전에 필터링하기 어려우며, 이메일스팸이 익명성 취약성을 갖는 것처럼 VoIP 스팸에서도 발신자 신분조작이 가능하다는 프로토콜 자체의 취약성이 있다.

[0008] 또한, VoIP 환경은 인터넷 이용이 가능한 환경이라면 스팸 발송이 가능하고, IP 주소 같은 논리적인 주소만으로는 지정학적 주소를 찾기 어려우므로, 스패머에 대한 실시간 추적이 실질적으로 불가능하다는 문제점도 있다.

발명의 내용

해결 하고자하는 과제

[0009] 본 발명은 상기와 같은 종래 기술을 극복하기 위하여 제안된 것으로서 스팸 차단 장치 및 그 방법에 있어, 튜링 테스트를 실행하여 오탐율을 최소화하는 튜링 테스트 기반 스팸 차단 장치 및 그 방법을 제공하는 것을 목적으로 한다.

과제 해결수단

[0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 스팸 차단 장치는, 발신호 요청을 하는 발신자에게 튜링 테스트의 질의 메시지를 전송하는 질의부; 발신자의 응답 메시지와 질의 메시지의 정답이 일치하는 지 비교하는 판단부; 및 판단부의 비교 결과를 근거로 발신자의 상태를 전이시키는 스팸 대응 정책부를 포함할 수 있다.

[0011] 또한 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 스팸 차단 방법은, 발신자의 상태를 판단하는 단계; 발신자의 상태가 블랙 상태인 경우에 튜링 테스트 수행을 지시하는 단계; 발신자에게 튜링 테스트의 질의 메시지를 전송하는 단계; 발신자의 응답 메시지와 질의 메시지의 정답이 일치하는지 비교하는 단계; 및 비교하는 단계의 결과를 근거로 발신자의 상태를 전이시키는 단계를 포함할 수 있다.

효 과

[0012] 본 발명에 따르면, 종래의 기술에서 스패머로 분류하는 판단의 오탐율을 낮추고, 발신 전용의 소프트웨어를 이용한 대량의 스팸 전송을 차단할 수 있다. 또한 튜링 테스트의 결과를 발신자에게 전송하여, 정상적인 발신자에게 자신이 스패머로 분류 되었는지를 인지하게 할 수 있다.

발명의 실시를 위한 구체적인 내용

[0013] 그레이리스트(greylist) 기반 필터링은 각 발신자(caller)의 스팸 정도를 일정한 기준에 따라 분류하여, 소정의

전화스팸(spam over internet telephony; 이하 SPIT 이라 한다) 레벨을 초과하는 발신자를 블랙리스트로 추가시키는 필터링 기법이다. 여기서, SPIT 레벨은 그레이리스트의 근간이 되는 각 발신자의 스팸 지수를 의미한다. SPIT 레벨은 발신자의 발신 패턴과 스팸 발신자의 발신 패턴과의 유사 정도를 가늠하는 척도이며 발신자를 스패머로 규정하는 절대적인 기준은 아니다.

[0014] 이하에서는, 그레이리스트 기반 필터링의 알고리즘 등에 대해 상술한다.

[0015] 도 1은 SPIT 레벨 결정 모델의 상태 전이도를 나타낸다.

[0016] 도 1을 참조하면, SPIT 레벨 결정 모델은 3가지 상태로 분류된다. 각 상태들은 발신자의 상태를 의미하는 것으로서, SPIT 레벨을 지정하기 어려운 미지 상태(Su, 610), 특정 속성 값들의 변화로 인한 정상적인 발신자로 분류된 그레이 상태(Sg, 620)와 완전히 스팸 발신자로 분류된 블랙 상태(Sb, 630)로 표현된다. 이러한 3가지 상태 정보는 발신자를 어떤 부류로 보아야 할지에 대한 정의로 활용되며, 본 SPIT 레벨 결정 모델은 이러한 상태 간의 상태 전이를 담당하는 함수로서 나타난다.

[0017] 미지 상태(610)는 정적인 속성만이 의미를 가지며, SPIT 레벨은 '0'이 할당된다. 정적 속성에 해당하는 발신자 아이디(caller ID), 과금정책(payment) 및 인증(authentication) 방법 등을 고려하여 SPIT에 대한 의심을 할 수 있는 상태이다.

[0018] 그레이 상태(620)는 호 수신자수, 호 지속시간, 호 트래픽 양, 발신 횟수(rate) 및 호 거부율(rejection rate)을 사용하여, 0 내지 9 사이의 값으로 표현되며, 이를 스팸지수 또는 SPIT 레벨로 나타낸다. 상기 속성들은 미지 상태(610)에서 그레이 상태(620)로의 상태전이(S62)에 관여하고, SPIT 레벨의 증가(S64)에 영향을 미치게 된다.

[0019] 블랙 상태(630)는 SPIT 레벨이 '10' 이 되었을 때를 의미하며(S66), 이 상태의 호는 무조건 스팸으로 처리된다. 이는 발신자의 다양한 정보를 종합해 볼 때 자명하게 스패머로 분류되어 블랙리스트에 추가되는 경우라고 할 수 있다.

[0020] 상기 SPIT 레벨 결정 모델에 따라, 일단 그레이 상태(620)로 분류된 발신자는 계속 관리 대상이 된다. 또한, 블랙 상태(630)의 발신자 역시 그레이 상태(620)로 이동하지 못한다. 단, 예외 사항으로써, 사용자에 의한 입력과 시간의 흐름에 의해, 상태 전이가 발생할 수 있다. 외부에서 특정 발신자를 반드시 블랙 상태(630)에서 미지 상태(610)로 옮겨야 한다는 요구가 있을 경우, 이를 예외 사항으로 보고, 허가하게 된다(S70).

[0021] 또한, 미지 상태(610)에서 블랙 상태(630)로의 상태 전이(S68)도 사용자의 요청에 의해 이루어질 수 있는 예외 사항 중 하나이다. 이와 같이 상태 및 상태 전이를 관리하는 목적은 발신자의 상태전이의 명시적 정의를 통해 차후 발생할 수 있는 혼란이나 오류를 막기 위함이다.

[0022] 도 2는 SPIT 레벨 결정 모델의 정의를 도시한 도면이다.

[0023] 도 2를 참조하면, 모델의 구성요소는 시간 속성을 표시하는 T, 외부 입력 집합을 정의하는 X, 상태 집합을 정의하는 Q 및 상태 전이함수 집합인 △로 구성된다. 시간 속성 T는 일정 시간이 지난 후 생길 수 있는 상태변화를 정의하기 위한 속성이다.

[0024] 외부 입력 X는 하기 표1과 같이 7가지로 구분된다. 이들 중 외부 요청을 제외한 6가지 사항은 스팸 판단을 위한 기준으로 활용되는 정보들이며, 상태 전이 혹은 그레이 상태에서의 SPIT 레벨 변화에 영향을 미친다. 초기의 외부 요청 입력은 사용자 혹은 알고리즘에 의해 외부에서 요청되는 상태 변화 요구로써, 이 입력은 그레이 상태를 거치지 않고 상태전이를 일으키는 예외 조건을 나타낸다.

표 1

[0025]

입력 X	설명
외부 요청	관리자 및 외부 프로그램 요청
$Call_{RR}$	호 거부율(Call Rejection Rate)
$Call_{NC}$	호 수신자수(Number of Call Receiver)
$Call_{D}$	호 지속시간(Call Duration)
$Call_{T}$	호 트래픽 양(Call Traffic)

$Call_R$	발신 횟수(Call Rate)
$Call_{IC}$	호간 휴지시간(Inter-Call Time)

[0026] 상기 호 거부율 내지 호간 휴지시간은 각각 미지 상태(610)에서 그레이 상태(620)로, 그레이 상태(620)에서 블랙 상태(630)로의 상태전이에 영향을 주는 입력이다. 또한, 그레이 상태(620)에서 내부의 SPIT 레벨을 계산하여 갱신해 주는 함수이기도 하다. SPIT 레벨이 1 과 9 사이인 경우, 계속 갱신은 되지만, 이로 인한 상태 변화는 발생하지 않는다. 모델의 상태는 현상의 대표 값을 정해 놓는 것이며, 본 모델의 정의에서 SPIT 레벨이 1 과 9 사이일 경우, 이를 그레이 상태로 정의했기 때문이다.

[0027] 도 1에서, 블랙 상태(630)에서 미지 상태(610)로의 상태 전이의 경우, 블랙리스트에 들어간 발신자를 일정 시간이 흐른 후 상태가 미지 상태로 바뀌는 상황을 고려한 것이다. 각 상태 전이 함수의 입력으로 상기 표 1에 명시된 발신자 관련 정보가 들어오면, 상태 전이 함수는 알고리즘에 의해 SPIT 레벨의 결정에 사용된다.

[0028] 다음으로 SPIT 레벨 결정 알고리즘을 고려한다. SPIT 레벨을 결정하기 위한 요소는 시뮬레이션을 통해 얻어진 데이터를 기반으로 한 것이다. 시뮬레이션에서 의미 있게 분석된 6개의 주요 데이터들은 전화 송수신을 통해서 얻어지는 정보들이며, 이 데이터들에 대한 특성 분석을 통해 SPIT 레벨을 결정하기 위한 알고리즘을 도출할 수 있다.

[0029] 하기 표 2는 상기 언급한 6가지 정보의 속성을 설명하는 표이다.

표 2

[0030]

	속성	설명
개별 호 속성이 고려되는 경우	호 수신자수	각 호의 수신자 정보를 기반으로 지표 도출
	호 지속시간	각 호의 호 지속시간 값을 사용하여 지표 도출
	호 트래픽 양	각 호의 평균 전송 트래픽량을 고려하여 지표 도출
여러 호 사이의 특성(통계)이 고려되는 경우	호 거부율	해당 발신자의 전체 호 중 거부된 비율을 표시
	발신 횟수	단위 시간당 해당 발신자의 호 생성 횟수를 표시
	호간 휴지시간	단위 시간당 해당 발신자의 호 사이 간격을 표시

[0031] 관리자가 파악하고자 하는 스팸 여부를 판단하기 위해 명시된 정보들이 수집되면, 각 정보들이 스팸과 관련성이 있는지를 분석하여, 도출된 값을 사용하여 해당 발신자가 스팸 발신자인지를 판단하면 된다. 각 지표들이 갖는 비중은 지표의 분류를 통해 초기 값을 정하지만, 스팸 발신자를 분류 관리하고자 하는 시스템에서는 이 비중 값을 수정해 가면서 관리할 수 있는 환경을 제공해야 한다.

[0032] SPIT 레벨을 결정하기 위해 필요한 속성들은 SPIT 레벨을 결정하기 위한 정량적인 값으로 도출되어야 한다. 특히, 각 지표에서 도출된 값을 적절히 정규화 하여 SPIT 레벨에 반영하는 것이 요구되며, 하기 표 3는 지표에 대한 계산식, 비중 및 값의 범위를 나타낸다. 값의 범위는 모든 지표에 대해서 0 과 1 사이의 값을 갖도록 설정하였다. 또한, 비중은 예시를 위해 임의로 설정된 것으로, 변경될 수 있는 값이다.

표 3

[0033]

지표	계산 방법	값 범위	비중 (초기값)
호 수신자수	전체 연결된 호 대비 호 수신자의 개수 계산 (Call recipient number/Call-Connected number)	[0...1]	50 %
호 지속시간	의미있는 통화로 볼 수 있는 시간 이상의 호 횟수를 연결된 호 횟수로 나누고, 역 비율을 계산	[0...1]	30 %

호 트래픽 양	현재의 평균 트래픽이 110% 이상의 평균 트래픽을 발생시키는 호에 대해, 스팸 호로 의심하고 전체 호 수에 대한 비율을 계산	[0...1]	10 %
호 거부 율	호 거부 율 값을 계산하고, 기존 데이터들의 분포를 활용하여 정규화	[0...1]	5 %
호간 휴지시간	호간 휴지시간 값을 계산하고, 기존 데이터들의 분포를 활용하여 정규화	[0...1]	3 %
발신 횟수	발신 횟수 값을 계산하고, 기존 데이터들의 분포를 활용하여 정규화	[0...1]	2 %

[0034] 호 사이의 관계(통계) 특성이 활용되어야 하는 경우(호 거부 율, 발신 횟수, 호간 휴지시간), 각 데이터의 평균과 표준 편차를 활용한 정규분포를 구성하고 이에 대해, 현재 데이터의 위치를 계산하는 방식으로 값을 결정한다.

[0035] 상기 호 거부 율, 발신 횟수, 호간 휴지시간 외의 3가지 지표, 즉, 호 수신자수, 호 지속시간 및 호 트래픽 양의 경우는, 상기 표 3의 과정을 따르며, 데이터의 특성을 보기 위한 정규분포를 만들어서, UI를 통해 제공한다.

[0036] SPIT 레벨 결정을 위해서 필요한 속성에 대한 결정이 끝나면, 각 속성들이 비정상적인 상태가 되는 상황에 대한 결정이 요구된다. 상기 결정을 위해서는 각 지표별로 임계값(threshold)이 필요하다. 임계값의 조절은 결국 블랙리스트에 해당 발신자를 포함시킬지 여부를 결정하는 실질적인 의사결정 기준이다.

[0037] 일정 수준 이상의 SPIT 레벨을 가지는 경우, 이를 SPIT 발신자로 결정하기 위한 결정 임계값(Decision Threshold)이 요구되며, 이 값은 0 에서 10 사이의 값을 갖는다. SPIT 레벨이 상기 결정 임계값 이상이 되는 경우, 해당하는 발신자를 블랙리스트에 포함시킨다. 블랙 상태로 분류된 발신자가 전송하는 호는 모두 스팸으로 처리된다.

[0038] 위의 알고리즘에 의하여 블랙 상태로 분류된 발신자는 위에서 언급한대로, 사용자에 의한 입력이나 시간의 흐름에 의해 상태전이가 될 수 있다. 이는 그레이리스트 기반 필터링의 알고리즘에 의한 상태 분류의 오류를 제거하기 위한 것이다. 하지만, 이 방법으로는 블랙 상태로 분류된 발신자 측에서는 자신의 상태를 전이시킬 수 없다. 또한 스패머로 분류되는 통화 패턴과 유사한 통화 패턴을 가진 정상적인 발신자를 블랙 상태로 분류하게 되는 문제점이 있다.

[0039] 따라서 이 문제점을 극복하기 위하여, 그레이리스트 기반 필터링에 튜링 테스트를 결합시킨 방법이 제안된다.

[0040] 그레이리스트 기반 필터링 중 블랙 상태로 분류가 되면 스패머로 인식되어 발신 호가 차단된다. 그러나 SPIT 레벨이 특정 발신자를 스패머로 규정하는 절대적인 기준이 되어 관리자가 설정한 임계값을 넘는 SPIT 레벨의 발신자들을 블랙 상태로 확정하여 이들의 발신호를 일괄 차단 한다면 블랙 상태의 통화 패턴과 유사한 통화 패턴을 가진 정상적인 발신자가 블랙 상태로 잘못 분류 될 수 있다. 따라서 그레이리스트 기반 필터링에 의해 스팸 발신자일 가능성이 높은 발신자에게 튜링 테스트를 시도하는 방법을 사용하여 상기와 같은 오류를 줄일 수 있다.

[0041] 도 3은 튜링 테스트 기반 스팸 차단 시스템의 구성도를 나타낸 도면이다.

[0042] 도 4는 본 발명에 따른 스팸 차단 장치의 구성을 나타낸 도면이다. 본 발명에 따른 스팸 차단 장치(400)는, 발신호 요청을 하는 발신자에게 튜링 테스트의 질의 메시지를 전송하는 질의부(410); 상기 발신자의 응답 메시지와 상기 질의 메시지의 정답이 일치하는 지 비교하는 판단부(420); 및 상기 판단부의 비교 결과를 근거로 발신자의 상태를 전이시키는 스팸 대응 정책부(430)를 포함할 수 있다. 또한 본 발명에 따른 스팸 차단 장치는, 상기 판단부의 비교 결과를 근거로 상기 발신자에게 발신호 요청에 대한 허용 여부를 알리는 메시지를 전송하는 인증부(미도시)를 더 포함할 수 있다.

[0043] 도 5는 본 발명에 따른 스팸 차단 방법의 순서도를 나타낸 도면이다. 본 발명에 따른 스팸 차단 방법은, 발신자의 상태를 판단하는 단계(S510);발신자의 상태가 블랙 상태인 경우에 튜링 테스트 수행을 지시하는 단계(S520); 발신자에게 튜링 테스트의 질의 메시지를 전송하는 단계(S530); 상기 발신자의 응답 메시지와 상기 질의 메시지의 정답이 일치하는지 비교하는 단계(S540); 및 상기 비교하는 단계의 결과를 근거로 발신자의 상태를 전이시키는 단계(S550)를 포함할 수 있다. 또한 본 발명에 따른 스팸 차단 방법은, 상기 비교하는 단계의 결과를 근거로 상기 발신자에게 발신호 요청에 대한 허용 여부를 알리는 메시지를 전송하는 단계를 더 포함할 수 있다.

[0044] 아래에서는 도 3 내지 도 5에 따른 발명을 상세히 설명하도록 한다.

[0045] 발신자 측에서 발신호 요청 메시지를 보내면, 스팸 대응 정책부는 발신호 요청 메시지를 수신하여 발신자의 상태를 판단한다. 발신자의 상태는 이미 그레이리스트 기반 필터링에 의하여 결정된 미지 상태, 그레이 상태 및 블랙 상태 중 어느 하나를 말한다. 튜링테스트에 의한 상태 전이를 실행하기 위해서, 그레이 상태와 블랙 상태 사이에 하나의 상태(예비 블랙 상태)를 더 추가할 수 있다. 예비 블랙 상태는 스팸 발신자로 분류되기 전에 거쳐야 되는 상태이다. 예비 블랙 상태에 있는 발신자는 스팸 발신자 후보가 될 수 있다. 기존의 그레이리스트 기반 필터링에서는, 발신자를 미지, 그레이, 블랙 상태 등 3개의 상태로 분류하고, 블랙 상태인 발신자의 발신 호를 일괄 차단함으로써 스팸을 차단하였다.

[0046] 그러나 기존의 필터링 과정에서는 위에서 언급한 몇 가지 주요데이터에 의해 발신자의 발신 유형을 분류하였기에, 실제로 정상적인 발신자라도 블랙 상태로 분류될 수 있었다. 따라서 정상적인 발신자를 스패머로 판단하는 오탐율을 낮추기 위하여, 블랙 상태의 발신자에게 튜링 테스트를 실시하여 그 결과에 따라서 정상적인 발신자 또는 스팸 발신자로 구분할 수 있다.

[0047] 발신자의 상태가 미지 상태 또는 그레이 상태라면 스팸 대응 정책부는 튜링 테스트 시행을 지시하지 않고, 정상적인 발신자로 판단하여 발신호 요청을 수락한다. 발신자의 상태가 예비 블랙 상태라면 스팸 대응 정책부는 튜링 테스트를 시행할 것을 질의부로 지시한다. 또한 발신자가 튜링 테스트를 통과하지 못한 상태라면, 이후의 발신 호에 대해서는 더 이상 튜링 테스트를 실시하지 않고 발신 호를 지속적으로 차단한다.

[0048] 튜링 테스트 시행 지시를 수신한 질의부는 데이터베이스로부터 튜링 테스트 질의 메시지를 생성하여 발신자에게 질의 메시지를 전송한다. 여기서 튜링 테스트 질의 메시지는 음성 메시지이거나 CAPTCHA(Completely Automated Public Turing test to tell Computers and Human Apart)에 의한 메시지일 수 있다. 튜링 테스트 질의 메시지를 수신한 발신자는 질의 메시지의 질문에 답한 응답 메시지를 다시 질의부로 전송한다.

[0049] 튜링 테스트를 사용하는 것은, 그레이리스트 기반 필터링에 의한 발신자 상태 분류를 보안하고자 하는 시도이다. 그레이리스트 기반 필터링에 튜링 테스트를 결합하게 되면, 스팸 오탐율을 낮출 수 있기 때문이다.

[0050] 발신자로부터 응답 메시지를 수신한 질의부는 판단부에게 튜링 테스트의 결과를 판단해 줄 것을 의뢰할 수 있다. 데이터베이스에는 질의 메시지와 그에 대응하는 정답이 저장되어 있을 수 있다. 판단부는 질의 메시지의 정답과 응답 메시지가 일치하는 지를 비교할 수 있다.

[0051] 여기서 정상적인 발신자가 응답한 경우라면 데이터베이스에 저장된 정답과 일치하는 응답 메시지를 전송할 것이다. 그러나 발신자가 스패머라면, 스패머는 보통 기계적인 방법으로 대량의 스팸 메시지를 발송하기 때문에 튜링 테스트에 응답 메시지를 전송할 수 없거나 정답과 일치하지 않는 응답 메시지를 전송할 것이다.

[0052] 판단부는 튜링 테스트의 결과를 질의부 또는 스팸 대응 정책부로 전송할 수 있다. 질의부는 판단부로부터 수신한 튜링 테스트의 결과를 스팸 대응 정책부로 전송할 수 있다. 튜링 테스트의 결과를 수신한 스팸 대응 정책부는 발신자에게 응답할 수 있다. 여기서, 응답은 발신을 허용하거나 차단하는 것일 수 있다. 또한, 튜링 테스트의 결과가 일치하는 경우에, 발신자는 스패머가 아닐 가능성이 높으므로, 발신자의 상태를 그레이 상태로 전이시킬 수 있다. 튜링 테스트의 결과가 일치하지 않는 경우에는, 발신자는 스패머일 가능성이 높으므로, 발신자의 상태를 블랙 상태로 전이시킬 수 있다.

[0053] 본 발명에 따른 튜링 테스트 기반 스팸 차단 시스템은 발신자에게 발신호 요청에 대한 허용 메시지 또는 불허 메시지를 전송하는 인증부를 구비할 수 있다. 인증부는 튜링 테스트의 결과에 따라서, 발신호 요청에 대한 허용 메시지 또는 불허 메시지를 발신자에게 전송한다. 이는 정상적인 발신자가 튜링 테스트에 제대로 응답하지 못한 경우에, 이를 알려줌으로써 정상적인 발신자에게 불이익을 주지 않기 위함이다.

[0054] 이상에서 설명한 본 발명은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 변경 및 변형이 가능하므로 전술한 실시예 및 첨부된 도면에 한정되는 것은 아니다.

도면의 간단한 설명

[0055] 도 1은 그레이리스트 기반 필터링에 따른 SPIT 레벨 결정 모델의 상태 전이도를 나타낸 도면이다.

[0056] 도 2는 그레이리스트 기반 필터링에 따른 SPIT 레벨 결정 모델의 정의를 나타낸 도면이다.

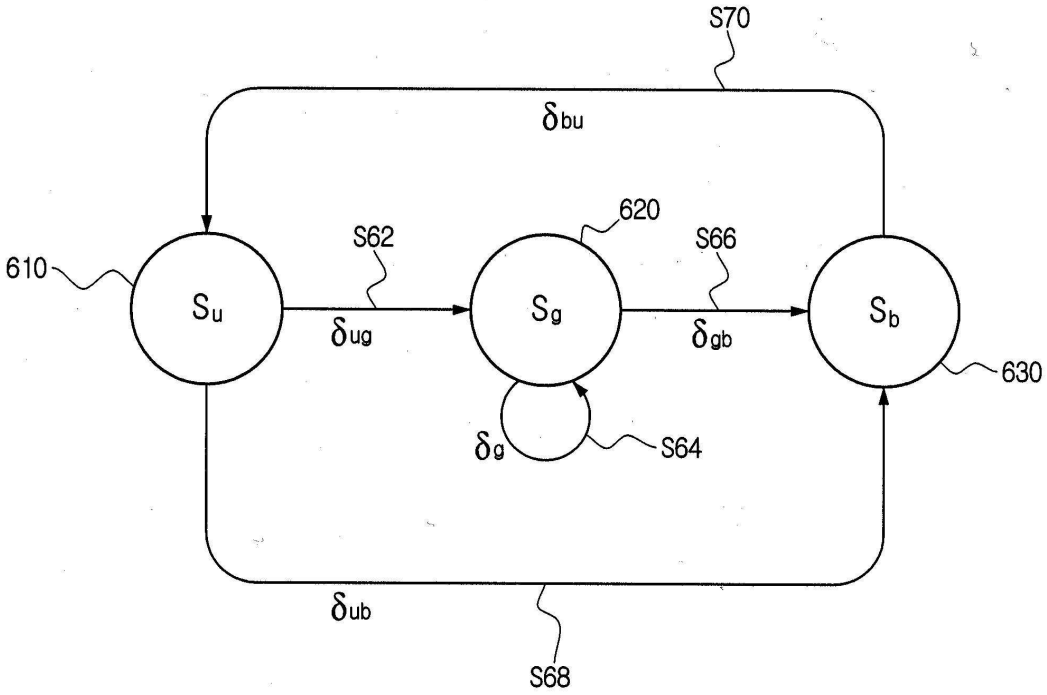
[0057] 도 3은 본 발명에 따른 튜링 테스트 기반 스팸 차단 시스템의 구성을 나타낸 도면이다.

[0058] 도 4는 본 발명에 따른 튜링 테스트 기반 스팸 차단 장치의 구성을 나타낸 도면이다.

[0059] 도 5는 본 발명에 따른 튜링 테스트 기반 스팸 차단 방법의 순서도를 나타낸 도면이다.

도면

도면1



도면2

SPIT Level Decision Model S=(T, X, Q, Δ)

Where

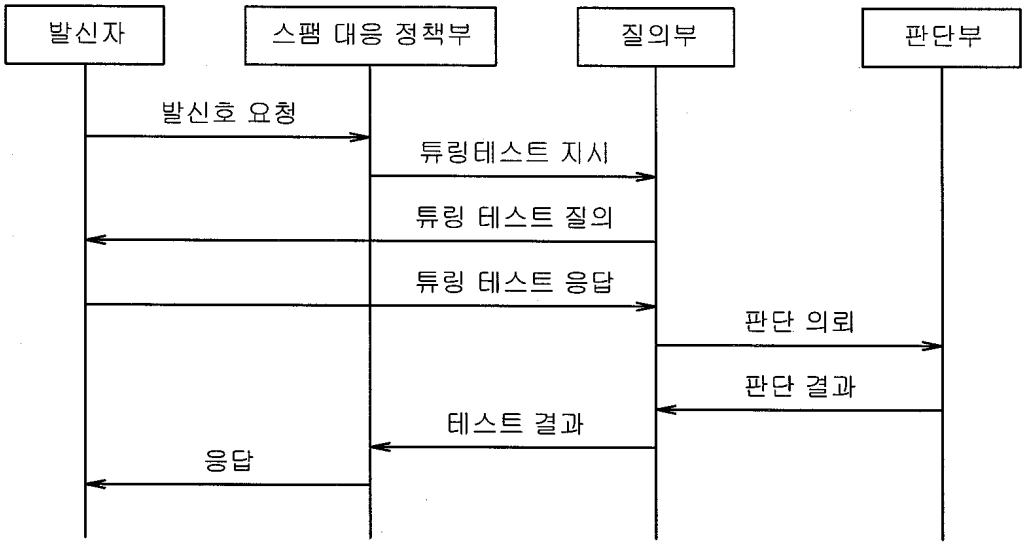
Time Element T

External Input X = {ExternalRequest, $Call_{RR}$, $Call_{NC}$, $Call_{D}$, $Call_{T}$, $Call_{R}$, $Call_{IC}$}

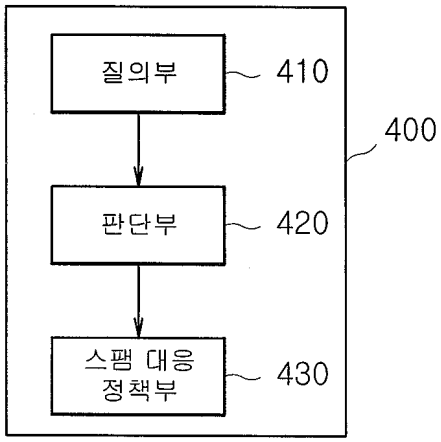
State Set Q = {s_u, s_g, s_b}

State Transition Function Δ = {δug, δg, δgb, δbu, δub}

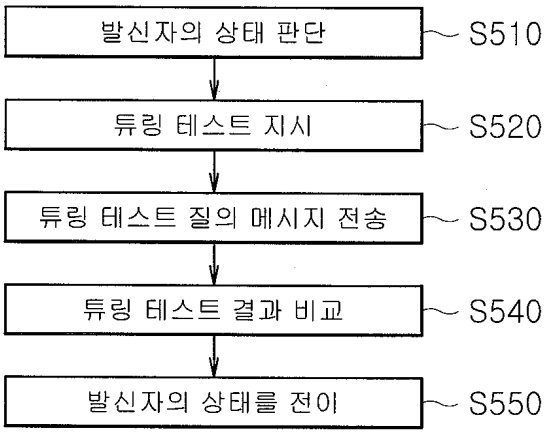
도면3


발신자
스팸 대응 정책부
질의부
판단부
발신호 요청
튜링테스트 지시
튜링 테스트 질의
튜링 테스트 응답
판단 의뢰
판단 결과
테스트 결과
응답

도면4


질의부
410
400
판단부
420
스팸 대응
정책부
430

도면5


발신자의 상태 판단
S510
튜링 테스트 지시
S520
튜링 테스트 질의 메시지 전송
S530
튜링 테스트 결과 비교
S540
발신자의 상태를 전이
S550