VoIP 스팸 Call의 Grey List 기반 SPIT 레벨 결정을 위한 정적 속성 분석 연구*

장은실** · 김형종*** · 강승석**** · 조영덕***** · 김명주****

요 약

VoIP 서비스는 사용자에게 기존 전화 서비스가 제공해 줄 수 없는 다양한 서비스를 제공해 준다는 장점 및 저렴한 가격으로 인해 사용자가 늘고 있는 추세이다. 다른 한편으로, VoIP 서비스의 저렴한 가격은 정상적인 사용자들에게만 매력적인 것이 아니라, 스팸 Call을 생성하는 사용자들에게도 유용한 환경이 될 수 있다. 본 연구는 스팸 Call을 탐지하기 위한 방법 중 그레이리스팅 기법을 사용하기 위해 정적·동적 속성을 분석 하였다. 또한, 정적 속성에 해당하는 인증 방법 및 과금 체계를 분석하기 위해서 국내의 VoIP 서비스 제공자의 서비스 제공 방식을 조사 분석 하였다. 이중, 시뮬레이션을 통해서 얻어진 Call 데이터를 사용하여, 스팸 Caller가 사용할 것으로 의심되는 과금 체계를 찾기 위해 데이터 분석을 수행하였다. 본 연구의 기여 점은 그레이리스팅 기법의 스팸 지수(SPIT Level)의 결정을 위한 VoIP Call의 정적 속성의 특성 분석에 있다.

Analysis on Static Characteristics for Greylist-based SPIT Level Decision of VoIP SPAM Calls

Eun-Shil Chang** · Hyoug-Jong Kim*** · Seung-Seok Kang****
Young-Duk Cho***** · Myuhng-Joo Kim****

ABSTRACT

VoIP service provides various functions that PSTN phone service hasn't been able to provide. Since it also has superiority in service charge, the number of user is increasing these days. When we think of the other side in cost aspect, the spam caller can also send his/her commercial message over phone line using more economic way. This paper presents the characteristics that should be considered to detect the spam call using greylisting method. We have explored static and dynamic characteristics in VoIP calls, and analyzed the relation among them. Especially, we have surveyed the authentication and charging method of Korean VoIP service provider. We have analyzed each charging method using our spam call simulation result, and derived the charging method that can be favored by spam caller. The contribution of the work is in analysis result of static aspect for SPIT Level calculation in greylisting method.

Key words : VoIP SPAM, SPIT, SPIT Level, Authentication, VoIP Payment

---

* 본 연구는 정보통신부 및 정보통신연구진흥원의 IT신신성장동력핵심기술개발사업의 일환으로 수행하였음 (2006-S-043-02, VoIP정보보호기술).

** 주저자, 서울여자대학교 대학원 컴퓨터학과
*** 교신저자, 서울여자대학교 컴퓨터학부 교수
**** 서울여자대학교 컴퓨터학부 교수
***** 한국정보보호진흥원 주임연구원

1. 서 론

VoIP(Voice Over Internet Protocol)는 PSTN과 모바일 통신에서 제공하던 음성 통신을 IP(Internet Protocol) 기반의 네트워크를 통해 제공되는 음성 통신을 말한다. 기존의 PSTN에서 IP로의 자원 사용의 확대와 비용 절감을 목표로 하는 서비스를 제공함으로써 초기 시장을 점유할 수 있었다. 그러나 안정적이고 일반화된 유선전화와 비교되는 통화 품질로 소비자의 만족을 이끌어내지 못하여 일반적인 통신 시스템을 대체할 수 있는 새로운 서비스 모델로 자리 잡지 못했다.

VoIP가 서비스되기 시작한 초창기는 소규모 서비스 업체들이 소비자가 요구하는 양질의 서비스를 제공하기 힘든 상황과 환경이었다. 그러나 지금은 기본 네트워크 인프라가 갖추어지고 IP 망의 안정화가 이루어졌으며 100Mbps 급의 광통신의 보급 등의 환경 변화로 저렴한 비용으로 이용할 수 있는 VoIP에 대한 관심이 급증하고 있다. 최근 몇 년 사이에 VoIP 서비스 업체가 급격히 증가한 것으로 사용자의 VoIP에 대한 수요의 변화를 볼 수 있다.

VoIP 서비스는 낮은 비용과 유선전화에서 제공하기 힘든 기타 부가 서비스를 제공하는 등의 이점을 가지고 있다. 기존의 유선전화를 대체하는 서비스로 전 세계에서 각광받고 있으며 통신 시장에서 급격한 성장을 보이고 있는 추세이다. VoIP 서비스는 이미 무시 할 수 없는 정도로 국내 시장의 한 부분을 점유하고 있으며 성장률은 계속 증가될 것으로 전망되고 있다[2].

이렇게 VoIP 서비스에 대한 시장 크기의 증가는 문제의 발생 가능성도 증가시키고 있다. 휴대전화와 이메일에서 문제가 되던 스팸 Call과 스팸 메시지의 문제가 VoIP에서도 나타나고 있다. VoIP 서비스의 장점인 저렴한 비용을 이용하여 다량의 스팸 Call을 발생시킬 수 있는 가능성이 있기 때문이다. 스팸 Call 문제는 사용자의 불편을 초래하고 사회적 윤리 문제까지 일으킬 수 있는 가능성을 보인다. 휴대전화와 이메일의 스팸관련 문제들은 사회적 법률과 여러 방어 기술로 사용자의 불편을 막을 수 있는 방법을 많이 발견했다. 이 방법과 기술들을 VoIP에서도 적용할 수 있을 것이다.

이메일 서비스는 수신자가 원하지 않는 스팸 메일을 받지 않도록 여러 가지 방법을 사용하고 있다. 가장 보편적인 방법은 메일의 제목이나 내용을 가지고 필터링 하는 것이다. 그 외에도 메일의 근원지 정보의 진실성을 검증하기 위한 방법, 메일에 특정 정보를 기반으로 전자서명 기법을 사용하고 수신자로 하여금 공개키를 사용한 복호화를 통해 검증하는 방법, 메일 내용의 단어를 통계 수치에 적용하여 사용하는 방법 등이 있다. 이런 여러 가지 방법들은 이메일 스팸이 사용자에게 미치는 영향을 최대한 막는 것을 목표로 하고 있다.

스팸 Call은 VoIP 서비스가 상용화 되는 것에 큰 방해가 된다. 이메일은 실시간 데이터 송수신이 아니기 때문에 정상적 메일을 수신하는 것에 미치는 영향이 적지만, VoIP 경우 다량의 스팸 Call은 수신자의 정상적인 Call에 대한 수신을 방해 할 수 있다. 송신자와 수신자간의 단일 연결인 PSTN과 달리 VoIP는 개방된 IP망을 이용하여 다량의 스팸 Call을 발생시킬 가능성이 높다. 이미 VOIPSA(Voice Over IP Security Association)는 VoIP의 스팸 Call을 원하지 않는 합법적 콘텐츠로 분류하여 정의했다[3]. VoIP의 서비스에 영향을 미치는 스팸 Call의 심각성을 느끼고 대응방법을 구체적으로 찾기 위한 분석을 하고자 한다.

스팸 Call에 대한 판단은 스팸 Caller가 가지고 있는 스팸 수준(SPIT-SPAM for Internet Telephony)의 정보에 기반 한다. 정책으로 정해진 스팸 지수를 가지고 스팸 Call의 가능성을 알아볼 수 있다. 발신자가 어떤 스팸 지수를 가지고 있는지에 따라 수신자가 통신 연결을 결정할 수 있는 중요한 요인이 되기 때문에 스팸 지수의 정책결정은 매우 중요하다고 볼 수 있다. 이메일 스팸과 달리 스팸 Call은 연결된 이후에 스팸으로 결정되는 경

우가 많기 때문에 사전에 스팸 Call을 막기 위한 여러 가지 방법이 제시되고 있다. White List와 Black List에 여러 가지 방법으로 저장된 정보는 스팸 Call의 발생 가능성을 낮출 수 있다. 그러나 사전 정보가 주어지지 않은 상태에서 발신자의 Call이 이루어 질 때 White List와 Black List로 결정하는 것은 어렵다. 스팸 Call에 대한 판단이 어려운 발신자에 대하여 Grey List정책을 이용하여 사용자의 정보를 1차적으로 저장하고 이후의 Call에 대한 정보를 누적하여 스팸 Caller로써 판단을 결정하는 정책을 사용하는 것이 스팸 Call을 막는 하나의 방법으로 제시하고 있다.

스팸 Call로 결정되기 위한 스팸 지수는 정적 스팸지수와 동적 스팸지수가 있다. 본 논문은 발신자가 가지고 있는 정적 스팸 지수 중 비용부분에 대한 정책 책정에 관한 실험과 연구를 통해 스팸 지수의 연산에 적합한 영향을 줄 수 있는 값을 이끌어 내고자 한다. Call에 대한 스팸 가능성을 결정하는 Grey List 관리 정책 결정에 정적 스팸 지수 중 비용이 미치는 부분에 대한 분석연구를 수행하였다.

2. 관련 연구

“원치 않는 메시지의 다량 전송”이라 불리는 스팸은 이메일에만 국한되지 않고 사용자간의 통신이 가능한 어떤 시스템에도 악영향을 미칠 수 있다. SIP(Session Initiation Protocol-VoIP에 사용되는 프로토콜)는 음성, 영상, 인스턴트 메시지, presence 스팸을 포함하여 사용자간의 멀티미디어 통신을 위해 사용되므로 이메일처럼 스팸의 표적이 될 수 있다. 따라서 먼저 SIP 스팸을 이메일 스팸 특성과 비교한 후 논의된 이메일 스팸을 위한 다양한 가능성을 가진 솔루션을 조사해 볼 것이다. 그리고 이것을 SIP 스팸에 응용하는 방법을 모색해보고자 한다.

2.1 SIP 스팸의 종류

Call 스팸은 SIP의 INVITE 요청과 같은 세션 초기화 시도를 통하여 이루어진 상대방이 원하지 않는 대량의 Call을 말하는 것으로써 음성, 영상 또는 IM(Instant Message)이나 기타 통신의 세션을 연결하기 위한 시도를 말한다. 전통적인 텔레마케터가 만든 스팸이 SIP에 적용된 경우라고 볼 수 있다. IM 스팸은 원치 않는 다량의 인스턴트 메시지 집합으로서 이메일 스팸과 유사하다.

Presence 스팸은 수신자에게 IM 스팸을 보내거나 다른 종류의 통신을 야기하기 위하여 수신자의 친구목록(buddy list)나 White List에 발신자 명단을 추가 하는 방법으로 IM 스팸의 한 형태로 볼 수 있다.

2.2 SIP 스팸 대응 솔루션

콘텐츠 필터링(Contents Filtering)은 이메일의 내용을 미리 검열하는 기법으로서 이메일 스팸에는 유용하지만 Call 스팸에 적용하기에 부적합하다. 사용자가 Call의 내용을 보기 전에 필터링이 되어야 하지만 내용을 보지 않고 콘텐츠 필터링을 하는 것이 불가능하기 때문이다.

Black Lists는 스팸 발신자의 사용자 이름과 전체 도메인으로 구성된 주소들의 목록을 유지하여 스팸을 걸러내는 방법이다. White List는 Black List와 반대의 개념으로 사용자가 이메일을 받고자 하는 적법한 발신자들의 목록이다. 강력한 신분 인증 방식과 White List가 결합되면 스팸에 대항하는 좋은 방법이 된다. IM 스팸을 예방하는 방법에 SIP에서의 White List는 상당히 효과적이다. 이는 SIP가 친구목록을 제공하고 있으며, SIP가 이메일보다 신분 인증에 더 안전하기 때문이다.

동의기반 통신(Content-based Communi-cation)은 Black List나 White List와 함께 사용된다. 발신자 A가 수신자 B의 Black List나 White List에 존재하지 않는 상황에서 A가 B에게 대화를 시도

했을 때 A의 요청은 일단 거절되며 동의가 필요함을 알게 된다. 이어서 B가 A에게 접근하여 A가 통신을 시도했음을 알려주고 수신자 B는 A를 인가하여 White List나 Black List에 넣을 수 있다.

2.3 이메일과 SIP의 신분 인증

이메일 스팸 대응 메커니즘에서 있어 인증되거나 검증된 신분은 발신자의 정보를 증명하는 중요한 부분이다. 스팸 대응 기술 중 많은 부분은 White List나 Black List와 결합될 때 보다 더 효과적인데 이것은 강력한 형태의 신분을 요구한다. 강력하게 인증된 신분은 많은 안티 스팸 기술이 제대로 동작하도록 하는데 중요한 역할을 한다.

SIP는 각 도메인이 자신의 사용자들을 인증하는 것으로 발신자의 신분을 확인할 수 있는 해법을 제공한다. 도메인이 사용자의 신분을 인증하면 이 사용자로부터의 메시지가 다른 도메인에 배포되었을 때 발신 도메인은 발신자의 신분을 명확히 하는 단언을 할 수 있으며 이러한 단언(assertion)에 대한 검증을 위하여 서명을 포함시킬 수 있는 것이 SIP 신분 메커니즘이다.

3. Grey List 관리 정책

제 2장에서 관련 연구를 통해 이메일 스팸의 특성을 알아보고 VoIP 스팸에 적용할 수 있는 방법을 찾아보았다. 스팸 메일과 스팸 Call에서 가장 많이 쓰이는 방법은 White List와 Black List 기법이다. 한 번 이상의 Call이 이루어진 발신자의 Call 정보는 누적된다. 이 정보는 발신자의 SPIT(SPAM for Internet Telephony) 레벨을 나타내고 발신자가 White List 또는 Black List에 속할지를 나타낸다. 그러나 Call에 대한 누적된 정보를 가지고 있는 발신자가 아닌, 사전 정보가 없는 발신자의 Call이 발생했을 때 White List나 Black List로 초기 저장하기 위한 명확한 기준을 찾는 것은 어렵다. 이 상황에 적용할 수 있도록 스팸 Call에 대한 판단이 어려운 사전 정보가 없는 발신자를 Grey List에 우선 저장하는 정책을 서론에서 제시하였다. Grey List를 관리 하는 정책에 대해 알아보고 정책 결정에 필요한 SPIT 레벨 결정을 위한 지수 산정에 관한 값을 알아보도록 한다.

3.1 Grey List 관리 정책 요소

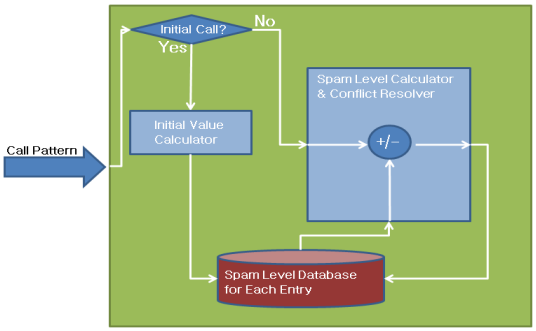
Grey List 관리 정책은 발신자의 동적·정적 정보를 기반으로 Grey List를 유지하기 위한 의사결정 규칙(Decision Making Rules)이라고 정의 할 수 있고 발신자의 정보 수집 모듈, 수집 정보의 사실화 모듈, 발신자의 SPIT 레벨 결정 모듈로 구성되어 있다.

발신자 정보 수집 모듈은 발신자의 Call과 관련한 raw data를 수집하는 역할을 한다. 여기서 수집되는 정보들은 Fact로 저장하기 위해 가공되기 직전의 상태로 준비 된다.

수집 정보의 사실화 모듈은 SPIT 레벨 결정 모듈이 처리해야할 정보를 생성해 준다. 생성되어야 할 정보는 Call Rate, Call Duration Consistency, Call Completion Success Rate, Cost of Call, Identity Strength의 5개 요소이고 Call의 통계적인 정보를 통해 사실화 되어야 한다.

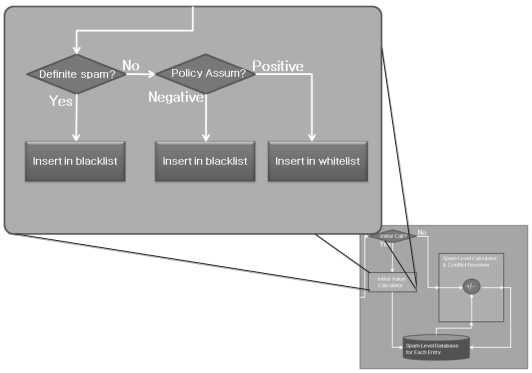
발신자의 SPIT 레벨 결정 모듈은 SPIT 레벨을 증가 또는 감소시키거나 초기 값을 설정해 주는 모듈이다. 해당 발신자의 Call을 수용할지 거절할지의 여부를 결정하는 데에 SPIT 레벨의 값을 활용하게 된다.

(그림 1)은 SPIT 레벨을 결정하기 위한 전체적인 논리 흐름을 보여주고 있다. 입력으로 제공되는 Call 패턴을 통해, 처음 시도된 Call의 경우 이의 초기 값을 설정해주고, 기존 값이 존재하는 경우 해당 엔트리를 찾아서 SPIT 레벨을 결정하기 위한 규칙을 새롭게 적용하여 나온 결과 값으로 기존의 SPIT 레벨을 조정한다.



(그림 1) Call 패턴을 스팸 수준 결정 전체 논통한 리 흐름

해결해야할 중요한 문제 중 하나는 사전 정보가 없는 발신자의 SPIT 레벨에 사용 될 스팸 지수를 어떻게 산출할 것이냐 하는 것이다. 사전 정보가 없는 경우에 대한 정책결정에 있어서는, Identity Strength와 Cost of Call과 같은 정적인 지수의 수집이 가능한 경우에 정책을 통해 표시해 줄 수 있다.



(그림 2) 초기 스팸지수 설정 방법

정보가 없는 Call의 경우 Black List에서 시작하는 것이 합리적일 것이다. 이는 사용자의 정책에 따라 특정 List로 가정하는 것을 시작으로 이에 대한 검증을 계속 하여 검증해 나가는 방법을 취해야 할 것이다.

예를 들어 사전 정보가 없는 발신자를 무조건 Black List나 White List에 넣은 후 해당 발신자의 누적행위를 분석 하여 그 위치를 옮겨가는 것이다. 초기 발신자의 Triplet 레코드를 Black List에 넣을지 White List에 넣을 지는 VoIP 운영관리자의 정책에 의해서 결정하는 것이 합리적이다. 그러나 해당 Call 정보가 자명한 스팸 Call의 형태를 갖고 있다면, (그림 2)의 예시처럼 초기 정보를 Black List에 넣는 것이 맞다.

3.2 스팸 지수 산정 요소

Grey List 관리 정책에서 SPIT 레벨 결정을 위한 5가지 요소는 발신자의 동적·정적 스팸 지수로 나눌 수 있다. 그 중 정적 스팸 지수는 Cost of Call, Identity Strength 두 가지로 나타낼 수 있다. 동적·정적 스팸 지수의 요소에 대해 알아보고 Grey List 관리 정책에서 어떻게 스팸 판정 기준으로 사용될지에 대해 분석해 보도록 한다.

정적 스팸 지수는 발신자의 신분(identity)의 인증 또는 Call에 사용되는 비용을 기반으로 한 특성이다. 이러한 특성은 Call의 주체인 발신자가 갖는 특성에 대한 정보와 함께 스팸 지수 산정에 대한 고려가 필요하다.

<표 1>에서 Identity Strength는 발신자의 신분을 인증한 방법에 대해 고려한 값이다. Level 0인 Unknown에서부터 Level 4의 “present passport”까지 강도의 표현이 나타나 있다. Cost of Call의 경우, Call 사용에 지불되는 비용의 정도를 의미하며, Level 0인 Unknown에서부터 Level 4의 “per Individually Call”까지 존재한다.

이러한 2개의 정적 스팸 지수에 대해서 상호 연관성이 고려되어야 하며, 이를 통해 정적인 Factor 기반의 SPIT Level을 도출해 낼 수 있다. <표 1>은 이러한 관계를 고려하여 정적인 속성을 명시적으로 알 수 있는 경우 SPIT 레벨의 결정에 활용될 수 있는 의심지수(suspect coefficient)를 제시하고 있다.

<표 1>의 요소 중에서 Identity Strength를 보면 Free, Paying Service, Physically Verified, Passport Presented로 나누었다. 여기서 Paying Service는 단순이 비용의 지불만 이루어지는 단계에서의 인증만 이루어지고 실제 사용 시에 개인 정보와 매핑 되는 것은 없는 방법을 말한다. Physically Verified는 물리적인 인증이 이루어지는 단계를 말하는 것으로 스마트카드나, S/W, H/W 적인 인증이 단계에 포함되어 사용자가 물리적 인증을 받기위한 단계가 필요하기 때문에 개인 정보가 어느 정도 매핑이 되는 것으로 발신자의 신용을 보장한다. Passport Presented는 실제 개인과 직접적으로 매핑이 되는 것을 표현하는 레벨로써 인증서나 개인 정보를 통해 검증받아 인증되는 것을 말한다.

〈표 1〉 정적 속성간의 상관관계를 고려한 스팸 의심지수

		Identity Strength				
		Unknown	Free Service	Paying Service	Physically Verified	Passport Presented
Cost	Unknown	N/A	N/A	N/A	N/A	N/A
	Free	N/A	1.0	0.75	0.5	0.25
	Flat Rate	N/A	0.75	0.5625	0.375	0.1875
	Per Minute	N/A	0.5	0.375	0.25	0.125
	Per Individual Call	N/A	0.25	0.25	0.125	0.0625

이렇게 Identity Strength의 스팸 지수 항목을 보면 Free 인증에서 스팸 지수가 가장 높다. 인증 단계가 많고 복잡할수록, 요구하는 개인 정보의 양이 많을수록 개인 정보와의 매핑 지수가 높아져 스팸 Call을 발신하는 사용자에게 개인정보 노출이 부담되기 때문에 스팸 발신 가능성이 낮아질 수 있는 근거로 작용할 수 있다.

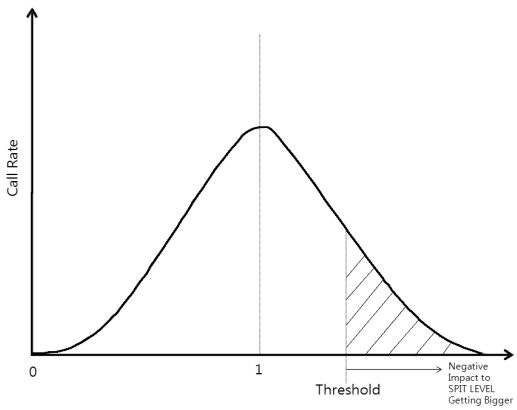
<표 1>에 나타난 정적 스팸 지수는 4×4의 행렬을 기준으로 최대값을 1로 하고 최소값을 1/16 (= 0.0625)으로 하여 계산되었다. Identity Strength 값을 1부터 4까지 값으로 정하고, Cost of Call을 1부터 4까지의 값으로 정하여, 이의 곱을 최대값 16으로 나누어서 0과 1사이의 값을 도출한 것이다. 이 값은 동적 특성을 고려하여 도출된 SPIT Level에 반영하기 위한 값으로 사용된다.

SPIT 레벨의 정책 결정 요소 중 정적 스팸 지수 한가지만으로는 스팸임을 결정하기 어렵다. 스팸 여부의 결정은 Call 흐름의 통계적 특성 및 통화 콘텐츠의 정보 등 동적 스팸 지수와 함께 이루어져야 한다. 스팸 여부 결정이 어려운 상황인 사전 정보가 없이 등장한 발신자의 경우, 초기 SPIT 레벨 값의 설정 시에 정적 스팸 의심지수가 활용될 수 있다. 즉, 신분 점검이 없는(free service) 발신자가 Identity Strength에 Cost of Call 값이 무료(Free)인 정적 지수를 나타낸다면 SPIT 레벨의 초기 값을 결정할 수 있는 근거로 사용할 수 있다.

정적 스팸 지수 외에 SPIT 레벨을 결정하는 동적 스팸 지수의 요소 중 하나인 Call Rate는 월 평균 대비 미리 지정된 단위 시간동안 시도된 총 Call 수를 의미한다. (그림 3)을 통해 나타난 Call Rate 값은 월간 평균 대비 값으로 값의 특성은 1을 평균값으로 하여, 0보다 큰 값을 갖는 정규분포의 형태로 비정상 특성을 고려할 수 있다. 이 값의 경우 1보다 큰 일정 임계치(Threshold)이상의 값을 가질 경우, SPIT 레벨의 상승에 크게 영향을 미치는 것으로 볼 수 있다. 그러나 1보다 작을 경우는 정상적인 상태로 인지하고 SPIT 레벨의 상승에는 영향을 주지 않게 된다.

$$CallCompletionSuccessRate = \frac{Num\_of\_CallSuccess}{Num\_of\_CallSuccess + Num\_of\_CallFail}$$

위의 수식은 Call Rate에 대한 값을 구하는 것



(그림 3) Call Rate의 정규분포

으로써 수식에 사용된 값인 Call Completion Success Rate는 실패한 Call과 성공한 Call의 비율을 표시한다. 이 비율의 분모는(성공한 Call 횟수 + 실패한 Call 횟수)이고, 분자는 성공한 Call 횟수이다. 이 값의 경우 1에 가까워질수록 SPIT 레벨에 음의 방향으로 영향을 미치고, 0에 가까워질수록 SPIT 레벨에 양의 방향으로 영향을 미칠 것이다.

$$CallDurationConsistencyFactor = |CallDurationGivenTime - AvgCallDurationformonth|$$

위의 수식은 Call Duration Consistency에 대한 값을 구하는 것으로 월 평균 대비 성공한 Call의 지속시간으로, 월평균 값에 대한 비율이다. 이것은 월 평균값과 비교하여 그 편차의 절대값을 고려하고 있다. 만일 편차 값이 일정 기준 이상으로 존재할 경우, 일반적이지 않은 경우가 발생한 것으로 인식할 수 있다. 즉 이 경우 편차가 일정 수준이상 큰 경우에 스팸 발생 문제로 여겨질 수 있기 때문에 이를 SPIT 레벨 결정에 반영하도록 한다.

<표 1>에서 나타낸 정적 스팸 지수와 수식으로 나타난 동적 스팸 지수를 연산하여 SPIT 레벨이 결정되고 사전 정보가 없는 발신자에 대한 Grey List 정책 결정에 사용된다.

4. 정적 스팸 지수 실제 사례 조사와 분석

제 3장에서는 정적 스팸 지수 에 대해 알아보았다. 이번 장에서는 정적 스팸 지수의 요소인 인증과 비용에 관한 실제 VoIP 서비스 업체들의 정보와 사례를 조사하고 분석하였다.

4.1 VoIP 서비스 업체의 인증 체계

사용자가 VoIP 서비스를 이용하기 위해서 기본 정보 이상의 인증을 받아야하기 때문에 발신자의 정보는 수신자에게 공개 될 수밖에 없다.

사용자의 인증 체계 요소는 발신자의 스팸 가능성을 나타내는 지수로 작용한다. 발신자가 어떤 인증 체계를 가지고 사용하느냐에 따라 SPIT 레벨의 산정에 있어 영향을 미치게 된다. 단순히 ID와 PW를 사용하는 인증 수준과 그보다 한 단계 이상을 추가한 여러 단계의 인증 수준을 가진 발신자의 스팸 Call에 대한 발생 가능성은 분명 다를 것이다.

4.2 VoIP 서비스 업체 과금 체계

정적 스팸 지수 중 비용인 과금 체계는 사용자의 인증 부분과 함께 사전 정보가 없는 발신자의 정체성을 잘 나타내는 부분이다. 어떤 요금제를 사용하고 있는지, 미리 저장된 정보에 따라 어느 정도 스팸 발신자로써 예상 할 수 있다.

VoIP를 서비스하는 업체들은 발신자가 사용 패턴에 따라 요금 체계를 결정할 수 있도록 다양한 요금제를 제공한다. 대표적인 요금제는 종량제와 정액제의 두 종류이다.

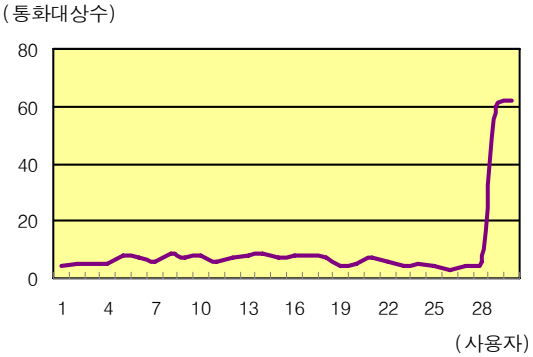
다량의 Call을 발생시키는 스팸 발신자에게 유리한 요금제는 정액제라고 할 수 있고 이것은 정적 스팸 지수를 산정하는데 있어 중요한 정보이다. 종량제는 사용하는 만큼의 요금을 지불하기 때문

에 발신 Call이 많은 스팸 발신자의 입장에서 불리한 조건이다. 정액제는 기본요금이 없이 정해진 요금을 지불한 후 사용 상한선을 정해놓고 사용하는 것으로 많은 양의 통화량을 제공하기에 스팸 발신자에게 더 유리하다고 할 수 있다. 이것은 스팸 발신자의 가능성이 종량제보다는 정액제 사용자에게 더 크다고 볼 수 있어 정적 스팸 지수 산정에 하나의 조건이 될 수 있다.

4.3 사용자들의 통화 패턴 기반 과금 내역 분석

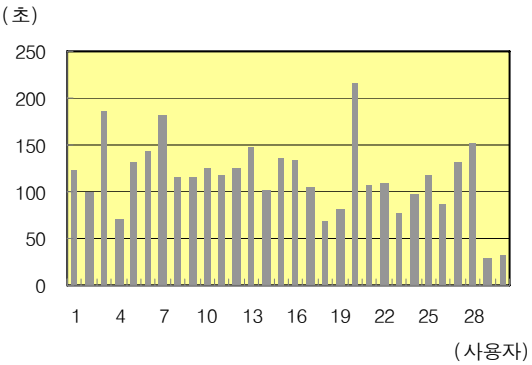
정적 스팸 지수의 하나인 요금은 동적이라고도 할 수 있다. 사용자가 어떻게 사용하는지의 패턴에 따라 요금 누적 상태가 다양한 상향 곡선을 그리기 때문이다. 실제로 정적 스팸 지수의 비용 관련 정보가 어떻게 변화하는 지를 분석하는 것은 스팸을 판정하는 데에 있어 의미 있는 부분이다. 실제 스팸 Call의 비용에 관한 정보를 분석하여 정적 스팸 지수를 산정하고자 실험하였다.

본 실험에서 사용된 Call 정보는 VoIP 시뮬레이션을 통해서 얻어진 데이터로, 스팸 Caller의 경우 짧은 시간에 많은 Call을 발생시키면서, 짧은 통화시간을 갖는 특성을 적용하여 사용하였다. 또, 스팸 Caller는 많은 Call 수신자에게 통화 시도를 하고 Silence 구간이 작은 이유로 트래픽 전송률이 높다는 것을 적용하여 실험을 하였다.



(그림 4) 사용자의 100명 당 통화대상수

(그림 4)에서 임의의 사용자 29와 30은 다른 사용자들과 통화 대상수에서 현격하게 많은 차이를 나타내고 있다. 이것은 스팸 Call이 통화 간격이 짧게 자주 일어나며 다양한 수신자를 대상으로 하기 때문에 통화대상수가 높을 수밖에 없는 것을 나타내고 있으므로 사용자 29와 30이 스팸 Caller의 가능성을 높게 나타낸다고 볼 수 있다.

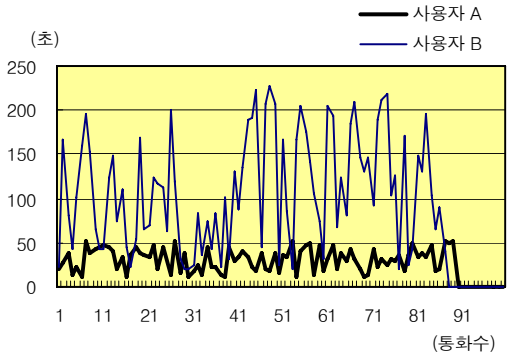


(그림 5) 평균 통화 시간

(그림 5)의 사용자 한 명당 100 통화에 대한 평균 통화시간 그래프에서 사용자 29와 30을 보면 평균 통화시간이 다른 사용자와 비교해 매우 짧은 것을 볼 수 있다. 스팸 Call이 걸려왔을 때 수신자가 스팸 Call임을 인지하는 순간 전화를 끊어버리는 경우가 많다. 이것은 스팸 Call의 통화시간이 짧게 유지될 수밖에 없음을 나타낸다.

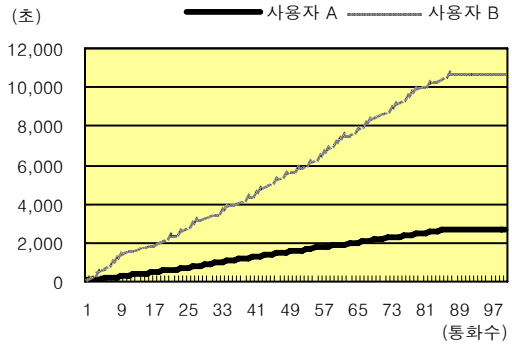
두 그래프의 비교를 통해 통화대상수와 평균 통화시간이 다른 사용자들과 차이가 많이 나는 스팸 Caller의 가능성이 있는 사용자 30과 일반적 통화 패턴을 보이고 있는 사용자 1을 비교 군으로 하여 사용자 30을 사용자 A로, 사용자 1은 사용자 B로 이름을 정하여 분석하였다.

(그림 6)의 그래프는 실험군인 사용자 A와 사용자 B의 100통화에 대한 통화시간을 나타내고 있다. 이 그래프를 통해 사용자 A와 사용자 B의 통화 패턴이 확연히 구분된다.



(그림 6) 통화 시간 분석

사용자 B는 짧은 통화에서 긴 통화까지 다양한 통화 시간을 가지고 있고 사용자 A는 비교적 짧은 1분 내외의 통화 시간대를 100통화 내에서 꾸준히 나타내고 있다. 이 통화패턴을 보았을 때 충분히 스팸 Call의 발생가능자는 사용자 A라고 할 수 있다. 사용자 A와 사용자 B에 대한 누적 통화 요금을 비교해 보면 그 결과는 더 명확히 드러난다.



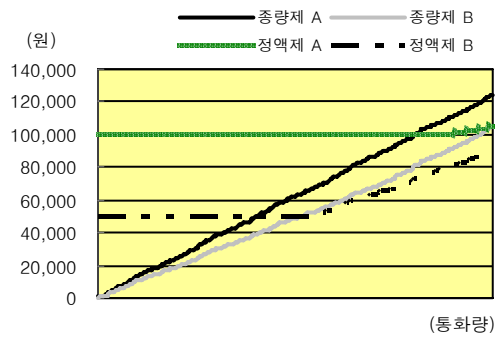
(그림 7) 누적 통화요금 분석

(그림 7)에서 100통화에 대한 요금을 시뮬레이션에서 실험값을 누적하여 3분(180)초당 요금을 37.6원이라고 정했을 때 사용자 A는 565.67원, 사용자 B는 2119.8원으로 계산된다. 이것은 약 4배의 차이를 보이는 것으로써 사용량이 100통화 이상 되었을 때 더 큰 요금 차이를 나타낼 것이다. 동일한 통화수의 사용자 중 통화 수에 대비해 유난히 적은 사용료를 내고 있는 사용자라면 스팸 Caller의 가능성이 있다.

그래프를 통한 결과를 보았을 때 평균 통화시간과 전체 통화시간들이 짧고 통화 수에 비해 적은 누적요금을 나타내고 있다면 스팸 Caller로 판단될 수 있다. 이 정보를 통해 정적 스팸 지수를 나타낼 수 있으며, 그 값은 SPIT 레벨 결정에 영향을 미치게 된다.

통화시간과 누적 요금의 스팸 지수 산정 가능성 이외에 VoIP 서비스의 대표적 요금제인 종량제와 정액제를 비교해 보고자 한다.

(그림 8)의 그래프를 보면 종량제 A와 B는 사용하는 만큼 요금을 지불하기 때문에 완만하고 꾸준한 요금의 증가를 보이고 있다. 종량제 A는 기본요금과 단말기 임대비용 이외에 추가의 초기 비용이 없이 사용하는 만큼 요금이 증가하고 종량제 B는 기본요금이 종량제 A보다 비싸고 단말기를 임대하는 것이 아닌 단말기를 구입 하는 등의 초기비용이 더 비싸기 때문에 통화 단위의 요금이 평균 더 싸게 책정되어 있어 종량제 A보다 상승지수가 낮게 나타난다.



(그림 8) 종량제와 정액제의 누적 통화 요금 비교

정액제는 요금제의 상한요금까지는 통화수에 상관없이 사용하다가 상한요금 이후부터는 종량제에 비해 저렴한 비용으로 통화료를 추가로 지불하게

된다. 정액제 A는 사용한도인 100,000원 이후의 요금 증가치가 사용한도가 50,000원인 정액제 B에 비해서 낮다. 정액제는 상한 요금이 큰 요금제일수록 그 이후의 요금 부과 체계가 더 낮도록 서비스를 제공하고 있기 때문이다.

이 그래프에서 종량제와 정액제를 비교했을 때 종량제의 상승비율은 정액제보다 크다는 것을 명확히 보여준다. 이것은 종량제의 요금 누적비율이 높다는 것으로써 통화량이 많을수록 요금 상승이 비율도 높아진다는 것을 의미한다. 또, 같은 요금을 지불했을 때 정액제에서 더 많은 통화가 가능하다는 것을 그래프를 통해 알 수 있다.

<표 2>에서 실제 서비스 업체인 A사의 두 요금제를 비교했을 때 동일한 요금을 지불한다면 정액제에서 저렴한 가격에 더 많은 통화를 할 수 있다. 이것은 정액제에서 비정상 Call인 스팸 Call을 발생할 가능성이 더 높다는 것을 보여준다.

〈표 2〉 종량제와 정액제 비교(A사)

	기본요금	요금상한	통화 수	통화시간
종량제	2,000원	39,600	약 2,632	약 131시간
정액제	39,600원	100,000	약 1,042	약 52시간

통화 시간과 요금에 관한 실험 결과를 가지고 정적 스팸 지수의 하나인 비용 부분에서 정액제를 사용하는 사람의 SPIT 레벨을 정할 때 더 높은 지수로 책정될 수 있다는 것을 보여주고 있다.

5. 결론 및 향후 연구

<표 3>에서 책정된 스팸 의심지수는 <표 1>의 정적 속성간의 상관관계 고려한 스팸 의심지수의 계산법을 토대로 산정된 스팸 의심지수와 Level이다. VoIP 서비스 업체들이 제공하는 요금과 인증 체계를 기반으로 스팸 의심 지수와 Level을 책정해 보았다.

무제한 정액제이면서도 비교적 간단하고 노출되기 쉬운 ID/PW 를 사용하는 인증 방법의 스팸 지수가 높게 나타하며 종량제를 사용하면서 H/W 인증을 통한 서비스를 사용할 때 스팸 지수가 가장 낮다.

〈표 3〉 요금과 인증에 관한 스팸 의심지수와 Level

			Identity Strength				
			Level 1	Level 2	Level 3	Level 4	Level 5
			Unknown	Free Service	ID/PW 인증	S/W 인증	H/W 인증
Cost	Level 1	Unknown	N/A	N/A	N/A	N/A	N/A
	Level 2	Free	N/A	1.0	0.75	0.5	0.25
	Level 3	무제한 정액제	N/A	0.75	0.5625	0.375	0.1875
	Level 4	정액제	N/A	0.5	0.375	0.25	0.125
	Level 5	종량제	N/A	0.25	0.25	0.125	0.0625

발신자의 기초 정보를 통해 스팸 지수를 산정하는 것에 그치지 않고 실제 책정된 동적·정적 스팸 지수들이 얼마나 유효한지 명확히 할 필요가 있다. 각 특성에 대한 신뢰성의 검증과 정확한 측정 결과를 통한 스팸 지수의 산정, 정적 스팸 지수에 대한 연산 요소의 신뢰성을 높일 수 있는 방법을 고려해야한다. 한 번 결정된 동적·정적인 지수는 SPIT 레벨의 연산에 큰 영향을 미치고 변경되기 어렵기 때문에 누구나 신뢰할 수 있는 방법을 찾아야 한다.

SPIT 레벨 결정에 영향을 미치는 정적 스팸 지수중 비용에 관한 실제 모델의 자세한 조사와 실험 분석을 통해 스팸 지수의 특성에 영향을 미치는 부분을 자세히 알 수 있었다.

향후, SPIT 레벨의 연산에 대한 연구가 진행 된

후 본 연구에서의 결론을 통해 정적 속성이 반영된 SPIT 레벨 연산 기법이 개발될 것이다.

참 고 문 헌

[1] 정수환, "VoIP 스팸과 보안", TTA Journal, Vol. 104J, 2005.

[2] 이종석, "Tech Guide-IP PBX 기반 VoIP -5. VoIP 시장 동향", 2006.

[3] VoIP Security and Privacy Threat Taxonomy, VOIPSA(VoIP Security alliance), 2005.

[4] Greylisting White Paper, http://projects.puremagic.com/greylisting/whitepaper.html.

[5] D. Richard Kuhn, Thomas J. Walsh, and Steffen Fries, "Security Considerations for Voice Over IP Systems", Technical Report, NIST, SP800-58.

[6] Voice Over Internet Protocol (VoIP) Security Technical Implementation Guide, Defense Information System Agency DISA, Jan 2004.

[7] J. Rosenberg, "The Extensible Markup Language (XML) Configuration Access Protocol(XCAP)", draft-ietf-simple-xcap-12 (work in progress), October 2006.

[8] J. Rosenberg, "Presence Authorization Rules", draft-ietf-simple-presence-rules-08 (work in progress), October 2006.

[9] J. Rosenberg, "A Framework for Consent-Based Communications in the Session Initiation Protocol (SIP)", draft-ietf-sip-consent-framework-01(work in progress), November 2006.

[10] G. Camarillo, "A Document Format for Requesting Consent", draft-ietf-sipping-consent-format-01(work in progress), November 2006.

[11] J. Rosenberg, "A Framework for Application Interaction in the Session Initiation Protocol (SIP)", draft-ietf-sipping-app-interaction-framework-05 (work in progress), July 2005.

[12] E. Burger and M. Dolly, "A Session Initiation Protocol (SIP) Event Package for Key Press Stimulus (KPML)", RFC 4730, November 2006.

[13] T. Hansen, "DomainKeys Identified Mail (DKIM) Service Overview", draft-ietf-dkim-overview-03 (work in progress), October 2006.

[14] J. Rosenberg, "Applying Loose Routing to Session Initiation Protocol (SIP) User Agents (UA)", draft-rosenberg-sip-ua-loose-route-00(work in progress), October 2006.

[15] J. Lyon, and M. Wong, "Sender ID : Authenticating E-Mai", RFC 4406, April 2006.

[16] J. Lyon, "Purported Responsible Address in E-Mail Messages", RFC 4407, April 2006.

장 은 실

2005년 서울여자대학교 정보통신공학부 멀티미디어 통신공학과 (공학사)

2007년~현재 서울여자대학교 컴퓨터학과 석사과정

관심분야 : 모바일 네트워크, 무선 네트워크, 센서 네트워크, 네트워크 보안, QoS, ad hoc 네트워크

김 형 종

1996년 성균관대학교 정보공학과(공학사)

1998년 성균관대학교 정보공학과(공학석사)

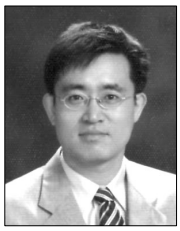
2001년 성균관대학교 전기전자 및 컴퓨터공학과 (공학박사)

2001년~2007년 한국정보보호진흥원 수석연구원

2004년~2006년 미국 카네기멜론대학 CyLab Visiting Scholar

2007년~현재 서울여자대학교 컴퓨터학부 전임강사

관심분야 : 취약점 분석 및 모델링, 이산사건 시뮬레이션 방법론, 침입감내기술



강 승 석

1992년 고려대학교 이과대학 전산과학과 학사
1998년 Michigan State University 전산학 석사
2004년 Michigan State University 전산학 박사
2005년 수원대학교 컴퓨터학과 전임강사
2006년~현재 서울여자대학교 컴퓨터학부 전임강사
관심분야 : ad hoc network, mobile computing, wireless communication, sensor network, QoS, anonymous communication, multimedia communication

조 영 덕

2000년 아주대학교 정보 및 컴퓨터공학부 졸업
2002년 아주대학교 정보통신 공학과 석사
2002년~현재 한국정보보호 진흥원 IT기반보호단 응용기술팀
관심분야 : VoIP 보안, 신종스팸 대응, 네트워크 보안, 신규IT서비스 보안

김 명 주

1986년 서울대학교 컴퓨터 공학과(공학사)
1988년 서울대학교 대학원 컴퓨터공학과(공학석사)
1993년 서울대학교 대학원 컴퓨터공학과(공학박사)
1993년~1995년 컴퓨터신기술 공동연구소 특별 연구원
2003년~2004년 미국 실바니아대학교(UPenn)객원 연구원
1995년~현재 서울여자대학교 정보보호학전공 교수
관심분야 : 정보보안, USN, 의료정보, 콘텐츠보안