Basit ve etkili bir spam engelleme yöntemi

Devrim Sipahi

devrim.sipahi@deu.edu.tr

Dokuz Eylül Üniversitesi

GİRİŞ

Spam, istenmeyen elektronik posta dır.

Bu tanım öznel bir tanım olsa da, spamların en az %90 herkes için spamdır.

Internetteki e-postaların %90 dan fazlası spamdır.

Bu durum ciddi bir zaman ve para kaybına yol açmaktadır.

Niçin Spam gönderiyorlar?

“Sayın Yetkili,

Firmanızın ve Ürünlerinizin tanıtımını size özel olarak hazırlanmış SMTP Mail Serverlar ile her ay milyonlarca kişiye toplu mail göndererek yapabilirsiniz. Aylık Sadece 600 TL Ödeyerek Milyonlarca Kişiye Ulaşabilirsiniz.

Günümüzde internet üzerinde yapılan ticaret hacminin ne kadar büyük olduğunu bilinmektedir ve buna bağlı olarak reklamlar da internet üzerinden yapılmaktadır. Toplu mail göndererek binlerce kişinin mail adresine ürünlerinizin bilgilerini gönderebilirsiniz. Size özel olarak hazırlanan bu Mail sunucularına uzak masa üstü bağlantısı yaparak sanki kendi bilgisayarınızmış gibi kullanabiliyorsunuz. %10 Spam toleransı bulunan sunucularımızın gönderimlerinin %90'ı inbox a düşmektedir. İsterseniz kendi mail portföyünüze, İsterseniz herhangi bir yerden temin ettiğiniz mail adreslerini kullanabilirsiniz.”

.....

Adresimizi Nasıl Öğrenmişler?

1. “Forward” (FW) e-postalar yardımıyla binlerce adres toplanabiliyor.

Önlem olarak yazacağınız adresleri Bcc kısmına yazabilirsiniz.

From: dr.atezi@windowslive.com
To: kavanozdancikanadam@hotmail.com; bogac_8181@hotmail.com; by_motobike@okubeni.com; capitalmail.tr@hotmail.com; esmerimmavi_52@hotmail.com; cevdet_altinisik@hotmail.com; isikhali.bepas@bayi.bellona.com; isikhali.bepas@bayi.bellona.com.tr; ceyda_and_ceyda@hotmail.com; ceydakarakullukcu@windowslive.com; cilemm16@hotmail.com; crazyy_ceylin@hotmail.com; dilaaaaa@hotmail.fr; dilar_helen777@hotmail.com; efe_ferhat6363@hotmail.com; gokom_35@hotmail.com; hasret-imsin@hotmail.de; herze-ipinhan@hotmail.com; hocakuruntusu@hotmail.com; ibrahimunner@okubeni.com; ibrahim_2021@hotmail.com; iyiinsanlargitti@hotmail.com; juni@ur.com; maricar4u20@live.com; merdan8647@hotmail.com; nurcanbaz11@hotmail.com; odemiskadir@hotmail.com; oftarik@msn.com; owner_xy@hotmail.com; sebo60@hotmail.com; dr.atezi@windowslive.com; selimsahinturk@hotmail.com; senem_uvz@hotmail.com; sessiz-tehlike@msn.com; seviyo61@hotmail.com; sibel_canan07@hotmail.com; sondurak_karatoprak_63@hotmail.com; vantu_tri@hotmail.com; zal_chee@hotmail.com; zekiustundag@hotmail.com; zeynep_y_8@hotmail.com

Adresimizi Nasıl Öğrenmişler?

2. Web sayfalarında text olarak yazılmış e-posta adresleri programlarla toplanabiliyor. .

Önlem olarak yazacağınız adresleri “@” simgesi kullanmadan veya resim olarak belirtebilirsiniz.

Örnek: devrim nokta sipahi () deu nokta edu nokta tr

Bu tür yazımları insan anlayabilir. Ancak programın anlaması zordur.

Adresimizi Nasıl Öğrenmişler?

3. Öğrenmelerine gerek yok. Kolaylıkla bulabilecekleri isimlere domain isimlerini ekleyerek adres tahmininde bulunabilirler.
Örnek ali@emo.org.tr devrim@deu.edu.tr

Öneri: Adresinizi bu şekilde tahmin etmişlerse sakın doğrulamayın.
Bu mailleri hiçbir şekilde yanıtlamayın.
Bu maillerdeki hiçbir linke tıklamayın.

Spam terimleri

Spam: istenmeyen e-posta
Ham: Spam olmayan e-posta
Positif <--> spam
Negatif <--> ham
False-Positif: spam olarak değerlendirilen
  ama gerçekte ham olan e-posta
False-Negatif: ham olarak değerlendirilen
  ama gerçekte spam olan e-posta

IP2: IPv4 adresinin ilk iki sekizlisi (oktet)

Spamların kökeni

E-posta sistemlerinin kullandığı protokol SMTP protokolüdür. RFC 821

Bu protokolde gönderen e-posta adresi ile gönderen IP adresi arasında herhangi bir denetim yoktur.

Dolayısıyla bir kişi gönderen kısmına istediği adresi ve domaini yazabilir.

Sender Policy Framework (SPF)

Bu protokol SMTP deki bu açığı gideren protokollerden biridir. Bu protokolde domain sorumluları, kendi domainleri adına e-posta göndermeye yetkili IP adreslerini DNS aracılığıyla yayımlar.

Alıcı taraf ise e-postanın gönderildiği IP adresinin DNS de yayımlanan domain adına yetkili IP ler içinde olup olmadığına bakarak e-postayı kabul veya reddeder.

Domain Name System (DNS)

DNS, Internetteki isimlerle IP adresleri arasında bağlantı kurar.
Örnek: www.deu.edu.tr <--> 193.140.151.17

Çeşitleri:

A kaydı: İsim -->IP

MX kaydı: E-postaları alacak isimleri söyler.

PTR kaydı: IP -->isim

TXT kaydı: isim --> bilgi

NS kaydı: ismi sorarsınız yetkili DNS makinesini söyler

Basit SPF kaydı nasıl yapılır?

Bunu için DNS teki TXT yapısı kullanılır.

Örnek: deu.edu.tr için SPF kaydı

BIND9 yazılımı:

deu.edu.tr. IN TXT "v=spf1 ip4:193.140.151.0/24 a mx -all"

A kaydı: 193.140.151.7

MX kaydı: 193.140.151.46 193.140.151.84

Yani: deu.edu.tr domaini adına e-posta göndermeye yetkili IP adresleri 193.140.151. ile başlamalıdır.

Ayrıntılar www.openspf.org ve RFC 4408 dedir.

Internetteki SPF kullanım oranı

SPF kaydı yapmak size gelen spamları engellemez. Sadece sizin adınıza sahte mailler gönderilmesini engeller. Eğer alıcı taraf SPF denetimi yapıyorsa. Bu nedenle kullanım oranı düşüktür.

5% Ekim 2006

9.9% spf-all.com Aralık 2008.

9.9% spf-all.com Ekim 2009.

Bu nedenle SPF tek başına yeterli değildir.

Sorular:

1. SPF kaydı **olmayan** domainlerin SPF kaydının nasıl olması gerektiğini tahmin edebilir miyiz?

Bu tahminin başarı oranını ölçmek uzun zaman alacaktır.

2. SPF kaydı **olan** domainlerin SPF kaydının nasıl olması gerektiğini tahmin edebilir miyiz?

Bu tahminin başarı oranını ölçmek için basit bir program ve çok sayıda domain yeterlidir. Yaklaşık (100,000)

SPF tahmin yönteminin temeli

Hemen hemen tüm domainlerin A ve MX kayıtları vardır.

Bazı domainlerin (%10) SPF kaydı (TXT) vardır..

Bu domainlerin A/MX kayıtları ile SPF kayıtlarındaki IP adresleri arasında ilişki vardır. Bu IP adresleri çoğu kez aynı network içinde yeralıyor.

Çünkü IP adresleri rasgele dağıtılmıyor.

Örnek:

193.0.0.0/8, 193.140.0.0/16 ve 193.140.150.0/24 IP adresleri sırasıyla RIPE, Ulaknet ve Dokuz Eylül üniversitenin sorumluluğundadır.

SPF tahmin yöntemi

Bir domainin SPF kaydı yok, ama A ve/veya MX kaydı var.

A kaydı: $a_1.b_1.c_1.d_1$ ve

MX kaydı $a_2.b_2.c_2.d_2$ olsun.

a, b, c, d sayıları 0-255 aralığındadır.

Bu durumda SPF kaydının aşğıdaki gibi olduğunu varsayıyoruz.
TXT: “v=spf1 ip4:$a_1.b_1$.0.0/16 ip4:$a_2.b_2$.0.0/16 -all“

Bir örnek

itu.edu.tr domaininin SPF kaydı yok. DNS bilgileri:

TXT: “ISTANBUL TECHNICAL UNIVERSITY”

A: “160.75.5.20 160.75.100.20 160.75.2.20”

MX: “160.75.2.5 160.75.2.2 160.75.2.3”

Yapılan SPF tahmini
TXT: “v=spf1 ip4:160.75.0.0/16 -all“

*@itu.edu.tr adresinen gönderilen bir e-postanın gönderen IP adresi “160.75.x.x” şeklinde ise spam değildir.

Farklı ise spamdır.

2. soruya yönelik örnekler

berkeley.edu A, MX ve SPF kaydı var.

A: “169.229.131.81” MX: “128.32.61.103”

Bizim tahminimiz:
TXT: “v=spf1 ip4:**169.229**.0.0/16 **128.32**.0.0/16 -all“

Gerçek SPF: “v=spf1 ip4:**128.32**.61.96/27 ip4:**169.229**.218.128/25

ip6:2607:F140:0:1000::/64 ~all”

Bizim tahminimiz gerçek SPF kaydını tamamıyla kapsamaktadır.

Diğer bir örnek

mit.edu A, MX ve Spf kaydı var.

A: “18.9.22.69” MX: “18.7.7.x 18.7.21.x”

Tahminimiz:
TXT: “v=spf1 ip4:**18.7**.0.0/16 **18.9**.0.0/16 -all“

Gerçek SPF: “v=spf1 ip4:**18.7**.7.0/24 ip4:**18.7**.21.0/24

ip4:18.72.0.0/16 ip4:**18.7**.68.0/24 ~all”

Bizim tahminimiz gerçek SPF kaydını kısmen kapsamaktadır..

SPF ve A/MX kayıtlarının karşılaştırılması

SPF kaydı olan domainlerin A, MX ve SPF kayıtlarındaki IP adreslerini bir programcık yardımıyla topladık.
Sonra bu IP adreslerinin ilk iki sekizliğini veritabanında bir tabloya aktardık.
A/MX kayıtları ile SPF kayıtlarını karşılaştırarak 3 grupta topladık.

1. A/MX kayıtları SPF kayıtlarının tamamını kapsıyor.
2. A/MX kayıtları SPF kayıtlarını kısmen kapsıyor.
3. A/MX kayıtları SPF kayıtlarını hiç kapsamıyor.

SPF ve A/MX kayıtlarının karşılaştırılması

Bu işlemi 3 farklı zamanda farklı sayıda domain ile yaptık.. Bu domainler, @deu.edu.tr adreslerine e-posta gönderenler arasında SPF kaydı olanlardan oluşmaktadır.

Karşılaştırma tarihi	Domain sayısı	Tam kapsama oranı (%)	Kısmen kapsama oranı (%)	Hiç kapsamama oranı (%)
Ocak 2007	2,074	51.70	41.00	7.30
Mayıs 2008	206,804	64.11	31.76	4.13
Ocak 2009	253,347	66.34	30.60	3.06

SPF ve A/MX kayıtlarının karşılaştırılması 2

Bu işlemi Kasım 2009 da spf-all.com sitesinden alınan 1 milyondan fazla domain için uyguladık.
Bu domainlerin yarısının spf kaydı yok.
Spf kaydı olanların %59 u “v=spf1 -all” şeklinde olduğu için değerlendirmeye alınmadı.
236513 tane domain karşılaştırmaya alındı.
C sınıfı temelinde yapılan karşılaştırmada tam kapsama oranı %64 çıktı.
B sınıfı temelinde yapılan karşılaştırmada tam kapsama oranı %70 çıktı.

SPF tahmin yönteminin uygulanması

Başlangıç tarihi Ocak 2007

Sunucu Sun V440

İşletim sistemi Solaris 10

mail sistemi qmail.

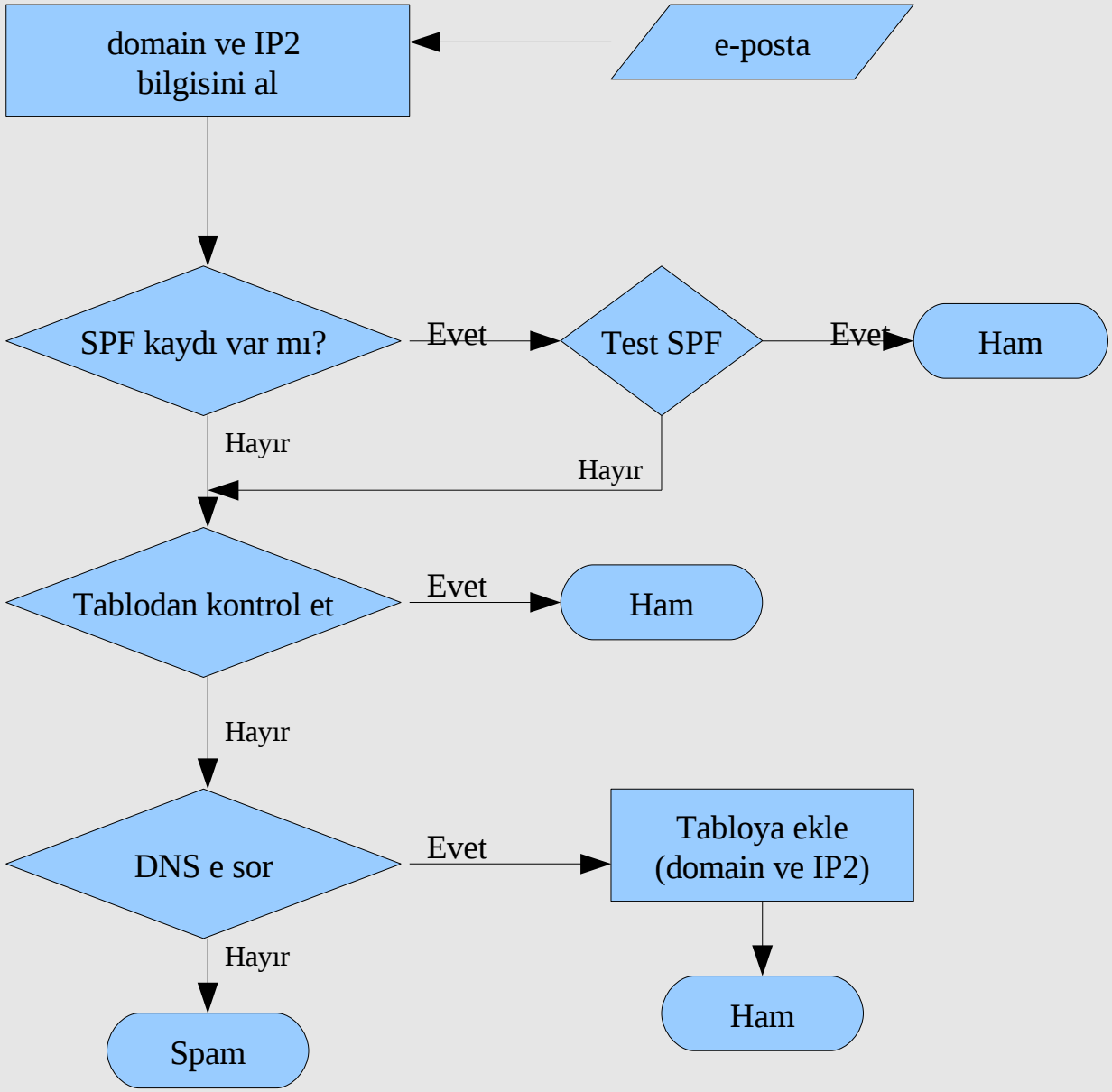
Program C dilinde yazıldı, ve .qmail aracılığıyla uygulandı.

DNS sorgularını azaltmak için domain ve IP2 (ilk 2 oktet) veritabanında (mysql) toplandı.

SPF tahmin yönteminin uygulanması

1. Gelen e-postalar önce kişinin karantinasına alınır.
2. Program çalıştırılıyor.
3. Spam olmayanlar kişinin postakutusuna aktarılıyor.
4. Spamların tarih, adres, konu ve IP bilgileri kullanıcıya bildirilmek üzere bir dosyada toplanıp günlük olarak bildiriliyor.
5. Spamlar siliniyor.

Programın Akış şeması



Spam istatistikleri

2007: Spam olmayan e-posta sayısı 1,670,926; ve spam sayısı 13,845,137 dir. (spam/toplam) %89

2008: Spam olmayan e-posta sayısı 4,276,611; ve spam sayısı 57,004,566 dir. Oran ise %93 tür.
Spam olmayanlar %156 artarken, spamlar %312 artmıştır.

Yani spamlar daha fazla artmıştır.

Geribildirimler

Kullanıcı memnuniyetini ölçmek amacıyla 3 anket yapıldı.

İlki Aralık 2007 de, katılımcı sayısı 510, toplam kullanıcı sayısı ie 1952 idi.

İkincisi Haziran 2008 de, katılımcı sayısı 684, toplam kullanıcı sayısı ise 3207 idi.

3. sü Kasım 2009 da yapıldı. Katılımcı sayısı 648, toplam kullanıcı sayısı 3833 idi.

Anket soruları

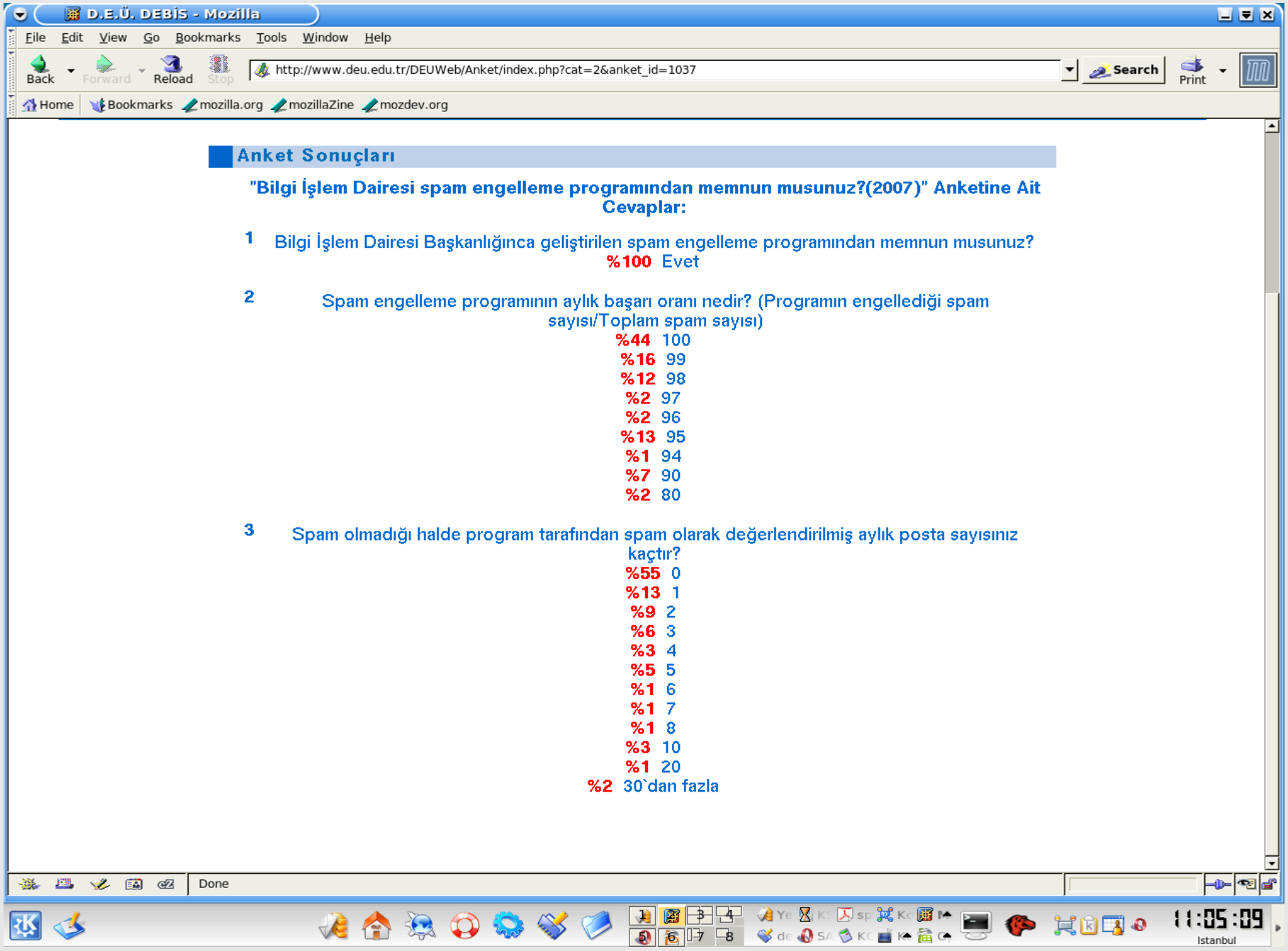
1. Programdan memnun musunuz?

2. Programın spamları engelleme oranı nedir? Yani gelen spamların yüzde kaçı program tarafından engellenmektedir. (engellenen_spam/toplam_spam)

3. Spam olmadığı halde program tarafından spam olarak görülen ortalama aylık e-posta sayısı kaçtır? (Aylık FP sayısı)

Anket yanıtları
(Ağırlıklı ortalama)

Tarih/ Sorular	Aralık 2007	Mayıs 2008
1. Memnun (Evet)	100.00%	100.00%
2. Spam engelleme oranı	97.35%	97,41%
3. Aylık FP sayısı	2.17	1.74



Anket Sonuçları

"Bilgi İşlem Dairesi spam engelleme programından memnun musunuz?(2007)" Anketine Ait Cevaplar:

1 Bilgi İşlem Dairesi Başkanlığınca geliştirilen spam engelleme programından memnun musunuz?

%100 Evet

2 Spam engelleme programının aylık başarı oranı nedir? (Programın engellediği spam sayısı/Toplam spam sayısı)

%44 100
%16 99
%12 98
%2 97
%2 96
%13 95
%1 94
%7 90
%2 80

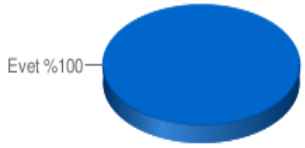
3 Spam olmadığı halde program tarafından spam olarak değerlendirilmiş aylık posta sayısınız kaçtır?

%55 0
%13 1
%9 2
%6 3
%3 4
%5 5
%1 6
%1 7
%1 8
%3 10
%1 20
%2 30`dan fazla

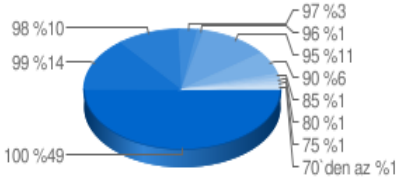
Anket Sonuçları

"Bilgi İşlem Dairesi Spam Engelleme Programı kullanıcı memnuniyeti anketi" Anketine Ait Cevaplar:

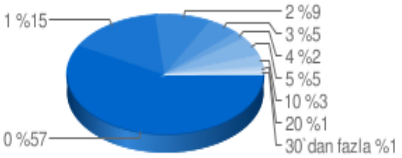
1 Bilgi İşlem Dairesi Başkanlığınca geliştirilen spam engelleme programından memnun musunuz? (683)



2 Spam engelleme programının aylık başarı oranı nedir? (Programın engellediği spam sayısı/Toplam spam sayısı) (684)



3 Spam olmadığı halde program tarafından spam olarak değerlendirilmiş aylık posta sayısınız kaçtır? (684)



D.E.Ü. DEBİS - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop http://www.deu.edu.tr/DEUWeb/Anket/index.php?cat=2&anket_id=1065 Search Print

Home Bookmarks mozilla.org mozillaZine mozdev.org

Anket Sonuçları

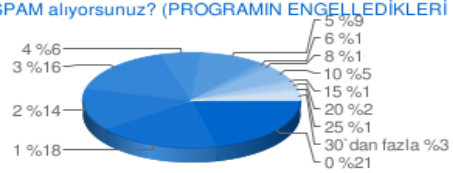
"Bilgi İşlem Dairesi Spam Engelleme Programı kullanıcı memnuniyeti anketi" Anketine Ait Cevaplar:

1

Kadro durumunuz nedir? (647)

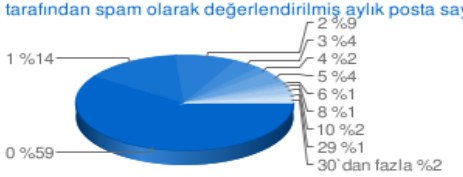
2

Günde ortalama kaç SPAM alıyorsunuz? (PROGRAMIN ENGELLEDİKLERİ HARİÇ) (648)



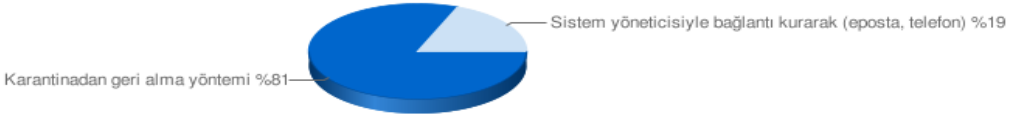
3

Spam olmadığı halde program tarafından spam olarak değerlendirilmiş aylık posta sayısınız kaçtır? (648)



4

3. soruda sayısını verdiğiniz spam olmayan e-postaları hangi yöntemi kullanarak geri alıyorsunuz? (407)



5

SPAM Engelleme programını ne kadar süredir kullanıyorsunuz? (648)

Done

Yerel Klasörler Konsole [3] Konqueror Kopete [2] Gimp [3]
KCalc Kwrite [5] Acroread [4] OpenOffice Mozilla-bin

17:27

Yöntemin zayıflıkları

1. Aynı networkteki iki kişi diğerinin domaini ile spam gönderebilir.
2. Yeni bir domain satın alarak spam gönderilebilir.
3. gmail, hotmail, yahoo gibi domainlerden bedava e-posta adresi alınarak az sayıda spam gönderilebilir.
4. Ele geçirilmiş veya yanlış ayarlanmış makineler üzerinden spam gönderilebilir.

Bunları önlemek için iki şey yapılabilir. İlki, bu adresleri (domain veya IP) karalisteye almak; ikincisi ise SpamAssassin gibi içerik tarayan ikinci bir filtre kullanmak.

SONUÇ

SPF tahmin yöntemi spam engellemek için iyi bir alternatif olmaya adaydır. Çünkü:
Spam engelleme oranı iyi.
False positif oranı düşük (1/5000)
Sadece e-posta başlığına baktığından hızlı ve az kaynak kullanmaktadır.
Ek bir donanıma ihtiyacı yoktur.

Bu başarı oranları spam gönderenlerin bu yöntemi henüz bilmediklerinden kaynaklanıyor olabilir.
Yöntemin gerçek başarısı ancak ilgili herkes öğrendikten sonra ortaya çıkacaktır.