SELÇUK
ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ

SUJEST, c.4, s.2, 2016
SUJEST, v.4, n.2, 2016
ISSN: 2147-9364 (Elektronik)

SOSYAL MÜHENDİSLİK İLE YAPILAN SALDIRILARININ DOĞAL DİL İŞLEME TEKNİKLERİ İLE ENGELLENMESİNE YÖNELİK WEB SERVİS GELİŞTİRİLMESİ

[1]Mustafa Ali AKCA

[1]*Süleyman Demirel Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, ISPARTA*

[1] mustafaakca@sdu.edu.tr

(Geliş/Received: 05.01.2016; Kabul/Accepted in Revised Form: 28.03.2016)

ÖZ: Siber korsanlar hedef sistemleri ele geçirme, geçici olarak servis dışı bırakma, tamamen kapatma, verileri silme gibi birçok farklı amaç için saldırılar düzenlemektedirler. Bu saldırılara çözüm olarak güvenlik duvarları, ağ izleme sistemleri vb. gibi birçok uygulamalar ile tedbirler alınmaya çalışılmaktadır. Günümüzde aslında belki bu saldırı yöntemlerinden çok daha etkili olan bir yöntem olarak kullanılan tekniklerden biride Sosyal Mühendisliktir. Sosyal Mühendislik; insan ilişkilerini ve insanların dikkatsizliklerini kullanarak, ikna etme, etkileme, aldatma gibi faktörlerle sistem, kurum ya da kişiler hakkında sıradan yöntemlerle elde edilemeyecek bilgilerin ele geçirilmesi işlemidir. Bilişim sistemlerinin güvenlik problemlerinde aslında korunması en zor olan yöntemlerden biride budur. Klasik saldırı yöntemleri çeşitli güvenlik yazılımları ile bertaraf edilirken burada saldırı yapanda hedef olanda insanın kendisidir. Bu çalışmada tüm masaüstü ya da web tabanlı uygulamalar için kullanılabilecek bir web servis geliştirilmiştir. Bu web servis Doğal Dil İşleme teknikleriyle karşıdaki kişiyi tanımaya çalışır. Karşıdaki aslında görüştüğü düşünülen kişi değil ise, farklı amaçları var ise kullanıcı uyarılır. Bu sayede sistemlerde Sosyal Mühendislik ile yapılan saldırılara da önlem alınmış olur. Çalışma site içi mesajlaşma ve email ile mesajlaşma uygulamalarında başarıyla test edilmiş ve olumlu sonuçlar alınmıştır.

Anahtar Kelimeler: *Sosyal Mühendislik, Siber Korsanlar, Doğal Dil İşleme*

Web Service Development For Restricting Attacks Done By Social Engineering With Natural Language Processing Techniques

ABSTRACT: Cyber hackers execute attacks with the many goal such as seizing target systems, leaving temporarily out of service, shutting down completely, and deleting data. In order to protect against those attacks, many applications such as firewalls and network monitoring systems are used. Today, in fact there is one technique that is much more effective than those hacking activities, which is Social Engineering. This term explains the action of seizing information about systems, institutions or individuals by using human relationships and people's carelessness, and persuasion, influence, and deception, which otherwise cannot be achieved by ordinary methods. This is actually one of the most difficult-to-protect methods in IT system security problems. While traditional hacking methods are eliminated by various security software, here the one who attacks and the one who is being attacked are both human itself. In this research, a web service which could be used both on desktop and web-based applications was developed. This web service tries to identify the opposite person by Natural Language Processing techniques. User is warned about it if the opposite person is actually not the person we think we are talking and has some different aims. Therefore, precaution could be taken against attacks done by Social Engineering. This service was tested on on-site messaging and e-mail applications and positives results were observed

DOI: 10.15317/Scitech.2016218523

Key Words: *Social Engineering, Cyber Hackers, Natural Language Processing*

GİRİŞ (INTRODUCTION)

Günümüzde teknoloji kullanımının artmasıyla birlikte kullanılan bu teknolojinin güvenliği konusunda her geçen gün daha da önem kazanmıştır. Bu da bilişim güvenliği kavramının ortaya çıkmasına neden olmuştur. Bilişim güvenliği denilince ilk akla gelen, sanal ortamda saklanan verilerin erişim izni olmayan kişiler tarafından erişilmesini önlemek, sistemlere üçüncü şahısların girmesini engellemek, gönderilen verilerin sorunsuz bir şekilde karşı tarafa gönderilmesini sağlamak gibi konuları içermektedir Günümüzde bir çok kurum, kendi içinde ve diğer kurumlarla olan haberleşmelerini internet tabanlı uygulamalar üzerinden yapmaktadırlar. Hem kurumsal hem de bireysel ihtiyaçlar için kullanılan yazılımlar sürekli olarak saldırılara hedef olmaktadır. Bu saldırı yöntemlerinden bazıları Cookie Hi-Jacking, ActiveX Saldırıları, TELNET saldırıları, FSO saldırıları, Hizmet aksatma saldırıları SQL injection saldırıları vb. olarak kategorilendirilebilir. Bu saldırı yöntemlerinde genel olarak amaç hedef sistemi tamamen ele geçirme, geçici ya da sürekli olarak hizmet dışı bırakmadır. Bu tür saldırılara güvenlik duvarları yazılımları kurularak, sunucu, ağ ve yazılım içinde yapılan iyileştirilmelerle önlemler alınmaya çalışılmaktadır. Günümüzde bu saldırılardan farklı olarak hedefin sistem değil insan olduğu bir saldırı türü vardır. Bu saldırı türü Sosyal Mühendisliktir.

Sosyal mühendislik ile yapılan saldırılar hedef kişiyi kandırarak, onun dikkatsizliklerinden yararlanarak normalde erişemeyeceği bilgilere erişerek sisteme ya da bireye zarar vermeyi amaçlayan saldırılardır. Günümüzde sosyal mühendislik yöntemiyle gerçekleştirilen saldırılardan bazıları aşağıda yer almaktadır:

Yetkili kişi aldatmacası: Bir firmadan hosting, domain ya da sunucu hizmeti alınmış olsun. Aslında firma ile sürekli haberleşilen yetkili bir firma emaili mevcut ve o firma emaili üzerinden sürekli haberleşme sağlanmakta. Ancak ilerleyen zamanlarda o firmanın adını taşıyan gmail ya da hotmail uzantılı bir emailden mesaj alınır. Bu email ile gönderilen mesajlar ilk zamanlar alınan hizmetle ilgili genel bilgiler içerir. Örneğin "Sunucularımızda bugün bakım yapılacak siteniz kısa süreli kesintiye uğrayabilir" ya da "Değerli müşterimiz sitenizin daha hızlı açılması için index.html dosyasını optimize edin". Bir süre bu tip emailler gelmeye devam eder. Bu sırada firmadan hizmet alan kişi o emaili firmanın yetkili emaili olarak kabul eder. Daha sonra kendisinden sistemle ilgili bazı bilgiler istenir. Örneğin : "Değerli müşterimiz, sitenizin config.php dosyasında bir hata tespit edildi. Hatanın düzeltilmesi için FTP kullanıcı adınızı ve şifrenizi lütfen gönderin". Birey uzun zamandır firma ile bu email üzerinden haberleştiğini sandığı için FTP kullanıcı adını ve şifresini yine bu email adresine yollar ve sosyal mühendislik tabanlı bu saldırıya düşmüş olur.

Arkadaş aldatmacası: Günümüzde en çok kullanılan sosyal mühendislik saldırı türlerinden biri de arkadaş aldatmacasıdır. Email üzerinden belirli aralıklarla yazışılan bir arkadaş isminden, bir süre sonra farklı bir email hesabından yine genel içerikli emailler göndermeye başlar. Örneğin "Bugün nasılsın, sınavların nasıl geçti". Hedef birey uzun bir süre bu email ile haberleşmeye devam eder. Daha sonra siber korsan bu bireyden acil bir durum olduğunu ve kendisine bir miktar göndermiş olduğu hesaba ödeme yapmasını ister. Hedef bireyde durumun farkında olmadığı için ödemeyi yapar.

Karşı hesapların ele geçirilmesi: Bu teknikte saldırganlar ele geçirdikleri bir facebook, twitter, email ya da site içi mesajlaşma uygulaması hesabıyla o bireyin tüm listesindeki kişilerle iletişim kurmaya devam ederler. Karşıdaki bireylerin hesabın ele geçirildiğinden haberi yoksa rutin olarak mesaj paylaşımlarına devam ederler. Bir süre sonra hesabı ele geçiren siber korsan, bireyin arkadaşlarından maddi bir menfaat talep edebilir. Karşıdaki bireylerin bu durumdan haberi yok ise ödeme yapabilir ve aldatılmış olurlar.

Sosyal mühendislik ile yapılan saldırılar tamamen karşıdaki bireyi aldatmaya (Can, 2008) yönelik olduğu için çok fazla yöntem ve teknik bulunabilir. Karşıdaki kişinin zaafları kullanılabilir. Onunla ilgili önceden tespit edilmiş bilgiler kullanılabilir. Örneğin "Ben uzun yıllardır görüşmediğin ilkokul arkadaşın". Kullanıcı bu aldatmacaya inanabilir. Sosyal mühendislik dijital saldırıların aksine aslında

günümüzde en çok zarar veren ve çözümü çok ta kolay olmayan bir saldırı çeşidi olmaya devam etmektedir.

Doğal dil işleme, ana işlevi doğal bir dili çözümleme, anlama, yorumlama ve üretme olan bilgisayar sistemlerinin tasarımını ve gerçekleştirilmesini konu alan bir bilim ve mühendislik alanıdır (Delibaş, 2008). Doğal dil işleme teknikleri ile dilin çeşitli özelliklerinden faydalanılarak birçok alanda çeşitli çalışmalar yapılmaktadır. Literatürde doğal dil işleme ile ilgili yapılan bazı çalışmalar genel olarak kategorilendirilmek istenirse "Yazar Tanıma" (Levent ve Diri, 2014; Yasdi ve Diri, 2012; Kaban ve Diri, 2008; Amasyalı ve Diri, 2006; Türkoğlu ve diğ., 2007; Diri ve Amasyalı, 2003),), "Metin Sınıflandırma" (Altınel ve diğ., 2013, Dural ve Diri, 2013; Çetin ve Amasyalı, 2013; Cingiz ve Diri, 2012; Biricik ve diğ., 2012),), "Soru Cevaplama" (Agan ve Diri, 2012; Topcu ve diğ., 2012;), "Metin Özetleme" (Güran ve diğ., 2014; Sami ve Diri, 2010; Uzundere ve diğ., 2008) vb. olarak gruplandırılabilir. Bireyler aslında herhangi bir konu hakkında metin tabanlı içerikler oluştururken (tez, makale, köşe yazısı, email, site içi mesaj, yorum, sms) bilinç altlarına işlemiş kendi karakterlerine ve kendi desenlerine uygun bir biçimde yazı yazarlar. Örneğin bazı bireyler devrik cümle kullanmayı çok sever, bazı bireyler noktalama işaretlerini hiç kullanmaz, bazı bireyler tüm metni büyük harfle yazmayı sever, bazı bireyler mesajlarının sonuna mutlaka "gülücük" eklerler. Bu ve buna benzer kişiye özgü birçok özellik vardır. Doğal dil işlemedeki bazı tekniklerde kişilere özgü olan bu özelliklerden faydalanarak içeriğe ait yazarın kim olduğunu tespit etmeye çalışırlar.

Bu çalışmada doğal dil işlemenin yazar tanıma teknikleri kullanılarak web servis destekli bir uygulama geliştirilmiştir. Bu uygulamaya email ve mesajlaşma uygulamaları ile bağlantı kurulabilmektedir. Bağlantı yapılan yazılımlar ilk olarak belirli bir süre bu yazılıma mesaj/email gönderen kişiyi ve mesajını metin olarak kaydetmektedir. Yazılım bu bilgileri belirli bir süre takip edip analizler gerçekleştirerek kişiye ait bir desen çıkarmaya çalışmaktadır. Mesaj gönderen kişiye ait yeterince içerik toplandıktan sonra ve analizler gerçekleştirildikten sonraki süreçlerde kişilerin göndermiş olduğu her mesaj önceki desenleri ile karşılaştırılmakta ve uyumsuz olduğunda "*mesajı alan kullanıcıya, karşıdaki kişinin aslında görüştüğünü düşündüğü kişi olmadığına*" dair bilgi verilmektedir.

MATERYAL VE METOT (MATERIALS AND METHOD)

Çalışmanın uygulama kısmında doğal dil işleme çalışmalarının yazar tanıma tekniklerinden bazıları kullanılarak, kişinin göndermiş olduğu metinleri analiz eden ve kişiye ait bir desen çıkarmaya çalışan bir motor yazılım yer almaktadır. Bu yazılım C# ile geliştirilmiş bir masaüstü uygulama olup Web Sunucusuna kurulmaktadır.

Şekil 1. Geliştirilen yazılımın çalışma prensibi

Figure 1. *The principle of the developed software*

Şekil 1'de çalışma prensibi görülen uygulama web servis desteği ile diğer uygulamalara açıktır. Herhangi bir email hizmetinden ya da herhangi bir site içi mesajlaşma hizmetinden uygulamaya bağlanılabilir ve veri gönderip analiz edilebilir. Uygulama analiz edilen kişinin bilgilerini saklayabilmek ve eş anlamlı kelimeleri karşılaştırabilmek amacıyla, eş anlamlı kelimelerinde bulunduğu bir MS SQL veri tabanına sahiptir. Uygulamanın sunucu üzerinde çalışan ana yazılımı hazırlandıktan sonra verileri test edebilmek amacıyla klasik ASP tabanlı bir site içi mesajlaşma uygulamasıyla, ASP.NET tabanlı POP3 desteğiyle email gönderen, posta kutusuna gelen emailleri okuyabilen bir web uygulaması geliştirilmiştir. Bu uygulamalar doğal dil işleme tabanlı yazılım ile web servis üzerinden iletişim kurabilmektedirler. Bir kullanıcı diğer kullanıcıya mesaj ya da email gönderdiğinde gönderen kullanıcının kim olduğu ve hangi mesajı yazdığı aynı zamanda bu yazılıma da gönderilmektedir. Yazılım yeterince bilgi topladıktan sonraki aşamalarda kullanıcıya bir mesaj ya da email geldiğinde kendi sayfasında kullanıcıyı bilgilendirebilecek bir uyarı mesajı göstermektedir. Bu uyarı mesajının yazılımdan alınması da yine bir web servis aracılığı ile yapılmaktadır.

Sunucuda Çalışan Yazılım (Software Runnig on the Server)

Doğal dil işleme ile yazar tanıma tekniklerinden biri de özellik (attribute) çıkarımı ile yazar tanımadır. Metin içerisinde belirlenen özellikler her alan, her konu, her şart ve durum için aynı olmayabilir. Özellik çıkarım işlemleri genel olarak her seferinde amaca uygun olarak optimize edilmektedir. Örneğin "Doğal Dil İşleme" ile yazar tanıma isimli bir çalışmada (Nesibe, 2013) 30 yazarın 40 köşe yazısından oluşan 1200 makalelik bir veri seti hazırlanmıştır. Bu veriler aşağıdaki özelliklere göre incelenmiştir;

- Ortalama paragraf uzunluğu (Boşluklu): Yazıda bulunan tüm paragrafların ortalama uzunluğu
- Ortalama paragraf uzunluğu (Boşluksuz): Yazıda bulunan ve boş olmayan paragrafların ortalama uzunluğu
- Ortalama cümle uzunluğu: Yazıdaki tüm cümlelerin uzunluklarının ortalama uzunluğu
- "Fakat" ve "Ancak" kelimelerinin sayısı: Yazıdaki "fakat" ve "ancak" kelimelerinin sayısı
- Boş paragraf yüzdesi: Boş paragraf sayısının tüm paragraf sayısına oranı
- Birinci tekil kişi sayısı: Birinci tekil kişi ile çekimlenmiş cümle sayısı
- Başlığın uzunluğu: Başlığın kaç karakterden oluştuğu
- Kaç adet sayı içerdiği: Yazıda kaç adet sayı bulunduğu
- Kaç adet paragraf içerdiği: Yazıda bulunan paragraf sayısı
- Özel isim yüzdesi: Yazıdaki özel isim sayısının tüm kelime sayısına oranı
- Yıldız işareti(*) yüzdesi: Yazıdaki yıldız işareti sayısının tüm noktalama işaretlerine oranı
- Nokta (.) yüzdesi: Yazıdaki nokta sayısının tüm noktalama işaretlerine oranı
- Virgül işareti(,) yüzdesi: Yazıdaki virgül işareti sayısının tüm noktalama işaretlerine oranı
- Başlıktaki noktalama işareti sayısı: Yazının başlığında bulunan noktalama işareti sayısı
- Çizgi (-) sayısı: Yazıdaki çizgi sayısının tüm noktalama işaretlerine oranı
- Çift tırnak işareti(") yüzdesi: Yazıdaki çift tırnak işareti sayısının tüm noktalama işaretlerine oranı
- Üç nokta işareti(…) yüzdesi: Yazıdaki üç nokta işareti sayısının tüm noktalama işaretlerine oranı
- Ünlem işareti(!) yüzdesi: Yazıdaki ünlem işareti sayısının tüm noktalama işaretlerine oranı
- Soru işareti(?) yüzdesi: Yazıdaki soru işareti sayısının tüm noktalama işaretlerine oranı
- Noktalama işaretleri yüzdesi: Yazıdaki noktalama işaretlerinin tüm harf ve sayıların toplamına oranı
- Noktalı virgül işareti(;) yüzdesi: Yazıdaki noktalı virgül işareti sayısının tüm noktalama işaretlerine oranı
- Stop-words yüzdesi: Yazıda bulunan alt başlık sayısının toplam kelime sayısına oranı
- Altbaşlık yüzdesi: Yazıda bulunan alt başlık sayısının toplam paragraf sayısına oranı
- Undefined word yüzdesi: Yazıdaki tanımlanmamış kelimelerin toplam kelime sayısına oranı

- "Neden", "Niçin" ve "Niye" kelimelerinin sayısı: Yazı içerisinde geçen "neden", "niçin" ve "niye" kelimelerinin sayısı
- Başlıktaki kelime sayısı: Yazının başlığında bulunan kelime sayısı
- Ortalama kelime uzunluğu: Yazıdaki kelimelerin uzunluğunun ortalaması
- Kelime uzunluklarının varyansı: Yazıdaki kelimelerin uzunluğunun varyansı (Nesibe, 2013)

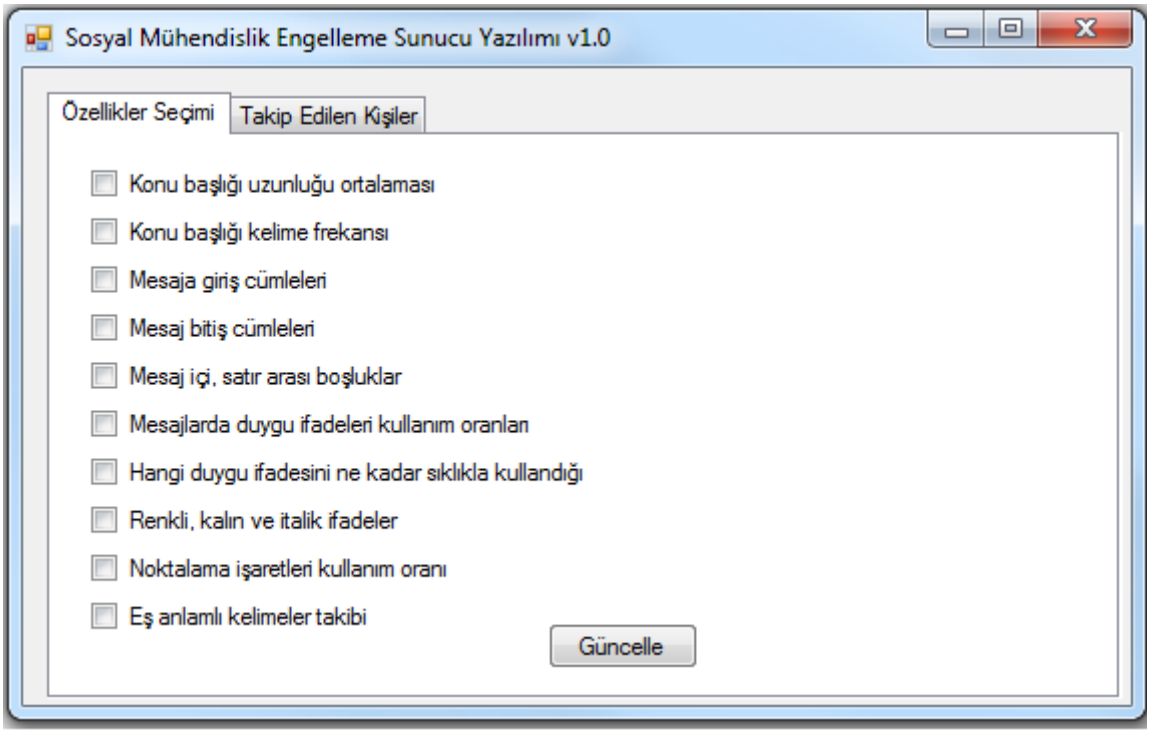
(Nesibe, 2013) yaptığı bu çalışmada verimli özelliklerin belirlenmesi ve her bir yazara ait çok fazla sayıda makalenin de olmasıyla %70 ile %80 oranında yazarı tespit eden sonuçlar elde edilmiştir.

Ancak sunucu için geliştirilen yazılım, çoğunlukla kısa metinleri analiz edeceği için bu özellikler tekrar gözden geçirilmiş ve verimlilikleri de test edilerek aşağıdaki şekilde çalışmaya uyarlanmıştır. Çalışmada kullanılan özellikler aşağıdaki gibidir;

- Konu başlığı uzunluğu ortalaması (Bazı bireyler içeriği her ne olursa olsun, tüm mesaj ve emaillerinde "selam", "merhaba", gibi ifadeler kullanırken bazı bireyler mesajlarının içeriğine uygun konu başlığı yazmaktadırlar)
- Konu başlığındaki "selam", "merhaba" gibi kelimelerin frekansları
- Mesaja giriş cümlelerinde sık kullanılan ifadeler. Örneğin her attığı mesajın ilk cümlesinde "Merhaba, nasılsın?" gibi ifadelerin kullanım oranı
- Mesaj sonunda, mesajı bitirdiği cümlede geçen kelimelerin sayısı. Örneğin : "tekrar görüşürüz", "kendine iyi bak"
- Mesaj içerisinde, satırlar arası boşluk bırakıp bırakmadığı. Enter tuşuna kaç defa basıldığı.
- Mesajlarda duygu (emotion) ifadelerinin kullanım oranları.
- Hangi duygu ifadesinin ne kadar sıklıkla kullandığı
- Renkli, kalın, italik ifadeler kullanıp, kullanmadığı
- Mesaj içerisinde virgül (,), üç nokta(...), soru işareti (?) gibi ifadelerin kullanıp/kullanmadığı, ne kadar sıklıkla kullandığı
- Eş anlamlı kelimelerden hangi kelimeyi daha çok kullandığı, ya da eş anlamlı diğer kelimeyi hiç kullanıp kullanmadığı (Türkçedeki eş anlamlı kelimeler yalın halleri ve ekleriyle birlikte veri tabanında bir tabloda tutulmuştur)

Makale, köşe yazısı vb. uzun metinlerde çıkarılabilecek bir çok özellik var iken içeriği kısa olan mesajlarda bu imkan oldukça kısıtlıdır ve başka özellikler tespit edilmesi gereklidir. Yapılan çalışmada yukarıdaki 10 özellik ile kişiler tanınmaya çalışılmıştır.

Ancak özellik sayısının az olması ve hata oranın yüksek olabileceğinin de düşünülmesiyle kullanıcıya verilen uyarı mesajlarının yanında "Hayır bu kullanıcı doğru sistemi eğitmeye devam et" özellikleri geliştirilmiştir. Bu sayede geliştirilen yazılım kişiyi tanıma adına sürekli olarak kendini eğitmeye devam edebilmektedir. Örneğin gülücük duygu ifadesi için birçok mesajında ":)" ifadesini kullanan bir kişi ":))" ifadesini kullandığında bu şüpheli bir durum gibi görünebilir. Ama bu uyarıyı gören kişi mesajda bir şüphe yok ise "Hayır bu kullanıcı doğru sistemi eğitmeye devam et" seçeneğini seçerek kişinin desenine bu duygu ifadesini de eklemiş olur. Gönderilen mesajlardaki metinlerin analizi için C# string sınıfından yararlanılmıştır. Her metin ilk olarak boşluklarına göre kelimelere ayrılmış, noktalama işaretler, kelime frekansları, enter tuşu kullanım frekansları tespit edilmiştir.



Şekil 2. Sunucu yazılımı arayüzü
Figure 2. *Server software interface*

Şekil 2'de görülen sunucu yazılımı arayüzünde takip edilen özelliklerin güncellenebileceği, takip edilen mesaj ve email kişilerinin ve hangi site üzerinden takip edildiğinin gözlemlenebileceği iki farklı sayfa bulunmaktadır. Bu yazılım web servis üzerinden birden fazla web uygulamasına destek verebildiği için her bir kullanıcı verisi site adresli olarak saklanmaktadır. Kişiye ait çıkarılan özellikler istenirse bu ara yüzden sıfırlanabilir. Bu sayede o kişiyi tanımak için gerekli olan eğitim tekrar başlamış olur.

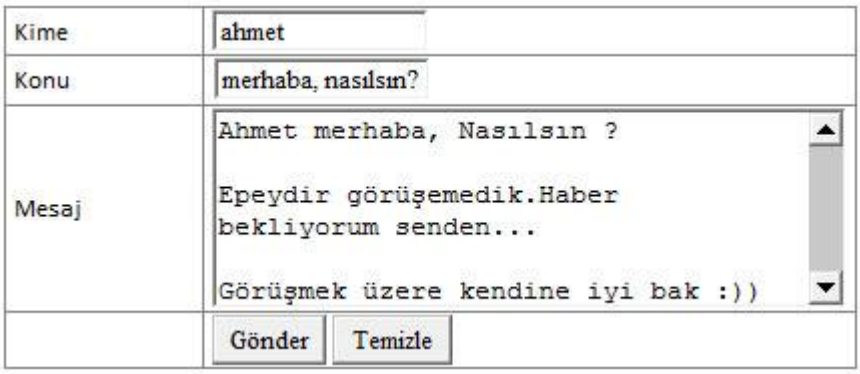
YAZILIMIN TEST EDİLMESİ (Software Testing)

Geliştirilen yazar tanıma yazılımı web servis desteği ile tüm uygulamalar üzerinden kullanılabilmektedir. Sistemin "Yazır tanıtma" ve "Yazar doğrulama" olmak üzere iki farklı webservis hizmeti yer almaktadır. Bu servisler aracılığı ile herhangi bir uygulama içinden erişilebilme imkanına sahiptir. Bu çalışmada site içi mesajlaşma ve email okuma/yollama uygulaması olmak üzere iki farklı uygulama üzerinden test edilmiştir.

Site İçi Mesajlaşma Uygulaması (Site Messaging Application)

Günümüzde facebook, twitter vb. başta olmak üzere birçok sosyal ağda site içi mesajlaşma uygulaması yer almaktadır. Geliştirilen yazılımın site içi mesajlaşma uygulaması ile entegreasyonunu test edebilmek için kullanıcı adı bazlı mesaj gönderen ASP tabanlı bir mesajlaşma uygulaması geliştirilmiştir. Uygulamaya üye olan kullanıcılar birbirleriyle site içerisinden mesajlaşabilmektedirler.

Şekil 3'te site içi mesajlaşma uygulamasının arayüzü görülmektedir. Uygulamaya üye olan kullanıcılar görünen ekrandan "kime", "konu" ve "mesaj" bilgilerini girerek mesajlarını karşıdaki kullanıcıya yollayabilmektedirler.



Şekil 3. Site içi mesajlaşma uygulaması arayüzü
Figure 3. Messaging application interface within the site

```
<%
' Bu kodlar bireyin karşıdaki kişiye gönderdiği mesajı
' veri tabanına kaydetmek için
set rs = server.createobject("ADODB.Recordset")
rs.open "SELECT * FROM MESAJLAR",baglanti,1,3
rs.addnew
rs("MESAJALICI")=request.form("kime")
rs("MESAJYOLLAYAN")=session("kullanici")
rs("KONU")=request.form("konu")
rs("MESAJ")=request.form("mesaj")
rs("TARIH")=now()
rs.update
rs.close : set rs = nothing

' Aşağıdaki kodlar mesajın bir kopyasını sunucudaki
' Doğal Dil İşleme ile Yazar tanıma yazılımına gönderen kodlar
sunucu = "http://localhost/ddi.aspx"
gonderensite = "http://localhost"
bilgi = "<Send><Gonderen>"& session("kullanici")
bilgi = bilgi & "</Gonderen><GonderenSite>"& gonderensite &"</GonderenSite><Mesaj>"
bilgi = bilgi & request.form("mesaj") &"</Mesaj></Send>"
Set xmlhttp = server.Createobject("MSXML2.ServerXMLHTTP")
xmlhttp.Open "POST", sunucu, false
xmlhttp.setRequestHeader "Content-Type", "text/xml"
xmlhttp.send bilgi
%>
```

Şekil 4. Site içi mesajlaşma uygulaması kodları
Figure 4. *Messaging application code within the site*

Şekil 4'te site içi mesajlaşma uygulamasının kodları görülmektedir. Kodların üst kısmında site içi mesajlaşma uygulamalarında veri tabanında mesajı alan/yollayan kullanıcıya göre saklayan kodlar yer almaktadır. Her mesaj "alıcı", "yollayan", "konu", "mesaj" ve "tarih" bilgileri ile veri tabanına kaydedilmektedir. Kodların alt bloğunda mesajın bir kopyasını sunucu üzerindeki doğal dil işleme yazılımına gönderen kodlar yer almaktadır. Mesajın kopyası gönderen kişi ve gönderen site bilgisi ile birlikte sunucudaki yazılıma gönderilir. Sunucudaki yazılım bu mesajı aldıktan sonra kişiyi tanımak için eğitilmek üzere mesajı kaydeder.

Mesaj gönderildikten sonraki karşıdaki birey kendi ekranında mesajı okuma sayfasında doğal dil işleme yazılımına bağlanıp yazar doğrulaması yapan bazı değişiklikler yapılmıştır. Şekil 5'te mesajı okuma arayüzü görülmektedir.

Kimden	Mehmet
Konu	merhaba, nasılsın?
Mesaj	Ahmet merhaba, Nasılsın ? Epeydir görüşemedik.Haber bekliyorum senden... Görüşmek üzere kendine iyi bak :))
Yazar Doğrulama Raporu	Gönderen Kişi Mehmet Değil Sebep 1: Mehmet şimdiye kadar göndermiş olduğu 78 mesajın hiç birinin girişinde "merhaba" içeren bir giriş yapmadı. Sebep 2 : Mehmet şimdiye kadar göndermiş olduğu 78 mesajın hiç birinde ":))" kullanmadı
İşlemler	Hayır bu kullanıcı doğru sistemi eğitmeye devam et Evet bu kullanıcı bilgisi şüpheli sistem eğitimine bu mesajı dahil etme
	Mesajı Sil

Şekil 5. Mesajı alan kullanıcı arayüzü
Figure 5. *User interface who take message*

Bu arayüzde standart mesaj uygulamalarındaki "kimden", "konu", "mesaj" bilgilerine ek olarak "yazar doğrulama" ve "işlemler" seçenekleride yer almaktadır. Bu ekranda kullanıcının göndermiş olduğu mesajdaki özelikler önceki mesajlarıyla karşılaştırılır. Önceki mesajlarındaki özellikleri ile yeni göndermiş olduğu mesajdaki özeliklerde farklılıklar var ise bu tespit edilip kullanıcı uyarılır. Örneğin Şekil 5'te kullanıcıya bazı uyarılar verilmiştir. Mesajı gönderen kişi önceki mesajlarından farklı bir mesaj girişi yapmış ve önceki mesajlarında şimdiye kadar hiç kullanmadığı bir duygu ifadesi kullanmıştır. Bu durum tespit edilmiş ve kullanıcıya bildirilmiştir. Bu gibi durumlarda her zaman şüpheli işlem olmayabilir. Kullanıcı ilk defa böyle bir giriş yapmış veya ilk defa böyle bir duygu ifadesi de kullanmış olabilir. Bu sebepten dolayı mesajı okuyan kişi bilgilendirilir ve ondan bir işlem yapması istenir. Mesajı gönderen kişiden şüphelenmiyorsa ve onun doğru kişi olduğundan emin ise "hayır bu kullanıcı doğru sistemi eğitmeye devam et" linki tıklanır. Bu sayede bu mesajda sisteme ilk aşamada farklı gelen özelliklerde mesaj gönderen kişiyle ilişkilendirilmiş olur. Ama eğer gerçekten şüpheli bir işlem var ise ve mesajı gönderen kişi doğru değil ise bu mesaj değerlendirmeye alınmaz.

Site İçi Mesajlaşma Uygulaması Testi (Site Messaging Application Test)

Çalışmadaki site içi mesajlaşma uygulamasının verimliliğini ve kullanılabilirliğini test edebilmek amacıyla 20 kişi ile 3 gün süren bir uygulama yapılmıştır. Her kullanıcıdan sisteme üye olduktan sonra gün içerisinde çeşitli kullanıcılara en az 20 mesaj atması istenmiştir. Mesaj içerikleri tamamen kişilere bırakılmıştır. Sistemde ilk 3 günde yaklaşık olarak 1.200 mesaj alışverişi yapılmıştır. Her bir mesaj göndericisine ait yaklaşık 60 mesaj kaydedilmiştir. Bu mesajlar kaydedilirken geliştirilen yazılım kişinin yazdığı mesajları, metin özellikleri çıkarımı yöntemiyle analiz edip kişileri tanımaya çalışmıştır. İlk 3 günlük sürenin sonunda, 20 kişilik grubun 10 kişisinin kendi aralarında kullanıcı adları ve şifrelerini karşılıklı değişmesi istenmiştir. Kullanıcı adları ve şifrelerini karşılıklı değişen kişiler ile kullanıcı adları ve şifrelerini kullanmaya devam eden grup 4. Günde kendi arasında mesajlaşmaya devam etmiştir. Testin gerçekleştiği 4. gün mesajlaşma işlemleri devam ederken, mesaj alan kişilerden 6 kişinin ekranlarında Sosyal Mühendislik saldırısı tespit edildiği uyarıları sık sık gözlemlenmiştir. Deney grubu olarak farklı hesaplardan 10 kişinin 6 tanesinin sahte kişiler olduğu başarılı bir şekilde gözlemlenmiştir.

Email Uygulaması (Email Application)

Çalışmadaki yazılımın email gönderme / alma uygulamaları üzerinden test edilebilmesi için POP3 üzerinden gelen emailleri okuyabilen / yeni email gönderebilen bir yazılım geliştirilmiştir. Bu yazılım C# tabanlı olup masaüstü uygulama şeklinde çalışmaktadır. Burdaki tüm teknikler tüm web tabanlı email gönderme/alma servislerinede entegre edilebilir. Geliştirilen yazılımın email gönderme sayfasında yukarıda yer alan mesajlaşma uygulamasındaki gibi email gönderildikten sonra aynı emailin bir kopyasını sunucu yazılımına web servis aracılığı ile gönderen kodlar yer almaktadır. Aynı zamanda bireylerin gelen emailleri okumuş olduğu sayfada da mesajlaşma uygulamasındaki ekrandaki bilgiler görüntülenmekte, email gönderen kişi, email konusu ve email içeriği dışında "Yazar Doğrulama Raporu" ve "İşlemler" seçenekleri eklemiştir. Bu uygulamanın tüm çalışma prensibi mesajlaşma uygulaması ile benzer şekildedir

SONUÇ ve TARTIŞMALAR (RESULTS and DISCUSSIONS)

Bilişim dünyasında internet kullanan tüm birey ve kurumlar her gün farklı saldırılara maruz kalabilmektedirler. Bu saldırıların birçoğu hizmet aksatma, hizmeti geçici veya sürekli engelleme, bilgileri ele geçirme, bilgileri değiştirme vb. gibi sistemlere, yazılımlara, web sitelerine yapıldığı gibi günümüzde bunlara ek olarak doğrudan bireyleri hedef alan "Sosyal Mühendislik" isimli saldırılarda popüler hale gelmiştir. Bu saldırı yönteminde karşıdaki bireyin dikkatsizliklerinden yararlanarak, karşıdaki bireyi aldatarak bireye zarar verme hedeflenmektedir. Diğer tüm siber korsan saldırı çeşitlerinde sistemleri savunabilen yazılımlar güvenlik duvarları kullanılabilirken "Sosyal Mühendislik" te yapılabilecek bir şey olmamaktadır. Çünkü bu doğrudan bireyi hedef alır.

Doğal dil işleme günümüzde metin özetleme, metin sınıflandırma, metin içerisinden duygu analizi, metni yazan yazarı tanıma vb. gibi birçok alanda kullanılabilmektedir. Doğal dil işlemede bu işlemleri yapabilmek için birçok farklı teknik bulunmaktadır. Bu çalışmada doğal dil işlemenin yazar tanıma tekniklerinden bazıları kullanılarak mesaj gönderen bireyler bir süre sistem tarafından takip edilmiş ve onlara ait bazı özellikler çıkarılmıştır. Bu sayede kişi sonraki mesajlarında, önceki mesajlarıyla karşılaştırılarak gerçekten mesaj yazan kişi o mu yoksa başka biri o hesabı ele geçirip mesajmı yazıyor tespit edilmesi hedeflenmiştir.

Çalışmada sunucu üzerinde çalışan web servis destekli yazar tanıma algoritmaları içeren bir yazılm yer almaktadır. Bu yazılıma ilk başta kişi ve kişiye ait mesajlar gönderilmiş, kişinin yazılım tarafından tanınması amaçlanmıştır. Mesajlar istenilen sayıya ulaştıktan sonraki aşamalarda ise, mesajı alan kullanıcıya, mesajı gönderen kullanıcı ile ilgili dönütler verilmiştir. Örneğin sürekli mesajlaşan iki kişiden birinin hesabı çalındığında ve diğeri bu durumdan haberdar olmadığında, mesajı alan kişi, mesajı gönderen kişinin sistem tarafından tanınmasıyla uyarılmıştır.

Çalışma site içi mesajlaşma ve email uygulamaları ile çok kez test edilmiştir. Site içi mesajlaşma uygulaması ile yapılan 20 kişilik ve 1.200 mesajlık bir testte %60 oranında başarı sağlanmıştır. Bu çalışmadaki yazar tanıma algoritmaları ve yöntemleri ihtiyaçlar doğrultusunda değiştirilebilir veya geliştirilebilir. Çalışma bu amaç doğrultusunda yapılacak sonraki çalışmalara model olma niteliğinde hazırlanmıştır. Bu çalışma ışığında site içi mesajlaşmalar ve email uygulamaları için yazar tanıyıcı sistemler geliştirilip, entegrasyon gerçekleştirilebilecektir. Sosyal mühendislik saldırılarını engelleme adına bundan sonraki çalışmalara da rehber olacağı düşünülmektedir.

KAYNAKLAR (REFERENCES)

Agan, C., Diri, B., "Aritmetik Bir Problemi Anlama ve Çözme: APAÇ", Akıllı Sistemler ve Yenilikler Uygulamaları, ASYU 2012, Trabzon, Türkiye, 2012

Altınel, B., Ganiz, M.C., Diri, B., "A Novel Higher-Order Semantic Kernel for Text Classification",10th International Conference on Electronics Computer and Computation, ICECCO 2013, Ankara, Turkey, 2013

Amasyalı, M.F, Diri, B., 2006, "Automatic Turkish Text Categorization in Terms of Author, Genre and Gender", 11th International Conference on Applications of Natural Language to Information Systems-NLDB2006, Austria, LNCS Volume 3999, Springer

Biricik, G., Diri, B., Sönmez, C., "Abstract Feature Extraction for Text Classification", Turkish Journal of Electrical Engineering and Computer Science, Vol.20, No.Sup.1, doi:10.3906/elk-1137-1159, 2012

Can, B., "Sosyal Mühendislik Saldırıları", https://www.bilgiguvenligi.gov.tr/sosyal-muhendislik/sosyal-muhendislik-saldirilari-3.html. Son Erişim : 28.12.2015

Cingiz, M. Ö, Diri, B.,"Mikroblog Kullanıcılarının Kategorizasyonu", IEEE 20. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU 2012, Fethiye, Türkiye, 2012

Çetin, M., Amasyalı, M.F., "Eğiticili ve Geleneksel Terim Ağırlıklandırma Yöntemleriyle Duygu Analizi", IEEE 21st Sinyal İşleme ve İletişim Uygulamaları, SIU 2013, KKTC, 2013

Delibaş, A., 2008, "Doğal Dil İşleme İle Türkçe Yazım Hatalarının Denetlenmesi", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, İstanbul.

Diri, B., Amasyalı, M. F., "Automatic Authorship Attribution in Turkish Language", 13th International Conference on Artificial Neural Network and 10th International Conference on Neural Information Processing, ICANN/ICONIP 2003, İstanbul, Turkey, 2003

Dural, B., Diri, B., "Türkçe Arama Motoru Kümeleme Sonucu Çalışmaları", IEEE 21st Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU 2013, Kıbrıs, 2013

Güran, A., Arslan, S.N., Kılıç, E., Diri, B., "Metin Özetleme için Cümle Seçim Metotları", IEEE 22.Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU 2014, Trabzon, 2014

Kaban, Z., Diri, B., "Yapay Bağışıklık Sistemleri ile Türkçe Metinlerde Tür ve Yazar Tanıma", 16. Sinyal İşleme ve Uygulama Kurultayı, SIU 2008, Aydın, Türkiye, 2008

Levent, V.E., Diri, B., "Türkçe Dokümanlarda Yapay Sinir Ağları ile Yazar Tanıma", Akademik Bilişim 2014, Mersin Üniversitesi, Mersin, 2014

Nesibe, Z., http://ziynetnesibe.com/dogal-dil-isleme-yazar-tanima-projesi., Son Erişim : 27.12.2015

Sami, M.V., Diri, B., "HTML Dokümanların Otomatik Özetlenmesi", Akıllı Sistemlerde Yenilikler ve Uygulamaları, ASYU, Kayseri, 2010

Topçu, S., Şen, C., Amasyalı, M.F., "Türkçe Sohbet Robotu", Akıllı Sistemler ve Uygulamaları, ASYU 2012, Trabzon,Türkiye, 2012

Türkoğlu, F., Diri, B., Amasyalı, M.F., "Author Attribution of Turkish Texts by Feature Mining", Third International Conference on Intelligent Computing, ICIC 2007, Qingdao, China, LNCS Volume 4681/2007 Springer, 2007

Uzundere, E., Dedja, E., Diri, B., Amasyalı, M.F., "Türkçe Haber Metinleri İçin Otomatik Özetleme", Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu, ASYU 2008, Isparta, Türkiye, 2008

Yasdi, M., Diri, B., "Soyut Özellik Çıkarımı İle Yazar Tanıma", IEEE 20. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU 2012, Fethiye, Türkiye, 2012