KULLANICI DAVRANIŞ ANALİZİ ile NÜFUZ TESPİT MODELİ (KDA-NTM)

Rahim KARABAĞ[1] **Hidayet TAKÇI**[2] **Türker AKYÜZ**[3]

Bilgisayar Mühendisliği Bölümü
[1,2,3]Gebze Yüksek Teknoloji Enstitüsü, PK. 141 41400 Gebze/Kocaeli

[1]rkarabag@gyte.edu.tr [2]htakci@bilmuh.gyte.edu.tr [3]takyuz@bilmuh.gyte.edu.tr

Anahtar Kelimeler: Nüfuz Tespit Sistemi,Bilgisayar Güvenliği, Kullanıcı Davranış Analizi, Veri Madenciliği ,k-NN

Abstract

Attacks on computer systems have been increasing as a result of attack tools developing and spreading rapidly and for this reason information security concept have become more important in all sectors. Most of the companies today understand the importance of information security and are using various protection methods. Technologies like firewalls, anti-virus software, intrusion detection systems, vulnerability scanners and encryption tools are all used in order to provide the security of information.

In this work, a novel method for Intrusion detection systems based on user behavior analysis has been proposed. User behaviors are obtained by the analysis of data packets on network. These are clustered by using data mining. After clustering process, users are inserted in to appropriate cluster. By the analysis of user data on network appropriate set of the user is found. When users behave different than these sets, an alarm will be given. As a result, intrusion detection by using user behavior analysis will build real time protection for network traffic.

Özet

Bilgisayar sistemlerine yönelik saldırılar, bu saldırıları gerçekleştirmek için kullanılan araçların hızla gelişmesi ve yaygınlaşması sonucunda artmakta ve bu nedenle bilgi güvenliği kavramı tüm sektörlerde bir hayli önem kazanmaktadır. Birçok kurum günümüzde, bilgi güvenliğinin önemini kavramış ve bu sebeple değişik koruma yöntemlerine başvurmuştur. Güvenlik duvarları, anti-virüs yazılımları, nüfuz tespit sistemleri, açıklık tarayıcıları ve şifreleme araçları gibi teknolojilerin tamamı bilginin güvenliğini sağlama amacına yöneliktir. Bu çalışmada, Nüfuz tespit sistemleri için kullanıcı davranış analizine dayanan bir yöntem önerilmiştir. Kullanıcı davranışları ağ üzerindeki veri paketlerinin analizi ile elde edilir. Kullanıcı Davranışları veri madenciliği kullanılarak kümeleme işlemine tabi tutulur. Bu kümeleme işlemlerinden sonra kullanıcılar kendilerine uygun kümelere dâhil edilirler. Ağ üzerindeki kullanıcı verileri analiz edilerek uygun küme bulunur. Bu kümelerin dışında davranışlar oluştuğu zaman alarm verilir. Sonuç olarak kullanıcı davranış analizi ile nüfuz tespiti yöntemi ağ trafiğinin gerçek zamanlı olarak korumasını sağlamış olacaktır.

1. Giriş

Saldırı, yetkisiz erişimlerle sistemin kırılmaya veya kaynakların yanlış kullanılmaya çalışılmasıdır. Saldırılar genellikle; hatalı ve eksik yetkilendirmelerde, zayıf şifreler kullanıldığında, sistem yanlış yapılandırıldığında ve yazılım kaynaklı açıklıklıklar bulunduğunda meydana gelir. Başarılı saldırı girişimleri nüfuz olarak tanımlanır. Saldırıların önlenmesinde kullanılan nüfuz tespit sistemleri güvenlik duvarının arkasında, ağ içerisinde çalışırlar [6,9].

Nüfuz tespiti için çeşitli yöntemler kullanılmakta olup bunlardan birisi de veri madenciliğidir. Veri madenciliği tabanlı nüfuz tespitinde daha çok sınıflandırma ve kümeleme gibi teknikler kullanılır. Eğer nüfuz tespiti için kümeleme seçilecek olursa o zaman önce ağ trafiği takip edilir ardından trafik durumuna göre kullanıcılar

bazı kümelerde gruplanır. Her bir grupta benzer trafik özellikleri gösteren kullanıcılar bulunur.

Bu çalışmada, nüfuz tespitinde, kullanıcı davranış analizine dayanan bir yöntem önerilmiştir. Yöntem, kullanıcıların iletmek üzere ağ üzerine yaydıkları veri paketlerinin analizini yaparak onları kümelere ayrıma esasına dayanır. Kümeleme sonrasında elde edilen kümelerin kimisinde normal erişim yapan kullanıcılar kimisinde ise saldırı girişiminde bulunan kullanıcılar yer alacaktır. Kümeleme sonrasında saldırgan olduğu kesin bilinen kullanıcılarla aynı kümede yer alan kullanıcıların da saldırgan olduğuna karar verilir. Buradaki varsayım benzer davranış gösterenlerin benzer trafik oluşturacağıdır. Trafik analizleri kullanıcı tabanlı yapılacaktır. Kümeleme tabanlı nüfuz tespiti anormallik tabanlı yaklaşıma benzer olup kümeleme algoritması için k-NN seçilmiştir.

2. Nüfuz Tespit Sistemleri:

Nüfuz tespiti, şüpheli faaliyetlerin tespitinde kullanılan metot ve tekniklerin genel adıdır. Nüfuz tespit sistemlerinde; saldırıyı başlamadan engelleme, yeni bir saldırı türü ise bunu öğrenip etkisiz hale getirme ilerisi için kural koyma, saldırganı yanıltma gibi teknikler kullanılabilir [6].

Nüfuz tespit sistemleri çalıştığı yere göre iki ana kategori altında incelenir:

1. *Konak(Host) tabanlı nüfuz tespit sistemleri:*Sadece kendi trafiğini dinleyen uç birimdeki bir yazılımdır.
2. *Ağ tabanlı nüfuz tespit sistemleri:* Ağ üzerindeki paketleri yakalayarak bunları çeşitli kıstaslara göre değerlendiren, ağ trafiği üzerinde yorumlar yapan sistemlerdir [9].

Nüfuz tespit sistemleri veri analiz yöntemine göre de iki kategoride toplanabilir.

1. *Kötüye Kullanım Tespiti:* Nüfuzları tanımak için daha önceden bilinen örüntülerden faydalanılır.
2. *Anormallik Tespiti:* Önce sistemin normal kullanım profili oluşturulur, ardından normalden sapmalar nüfuz olarak tespit edilir [5].

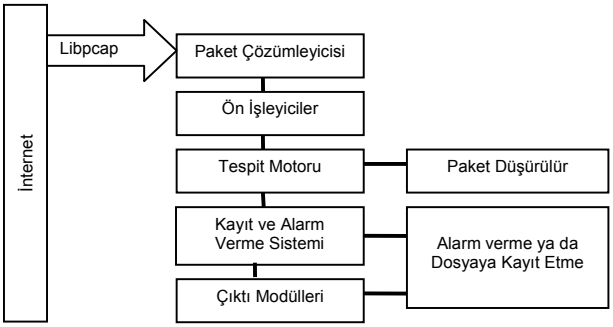
Kötüye kullanım tespitinde saldırı imzalarının tutulduğu saldırı veritabanları bulunur ve bu veritabanlarında yer alan imzalar elle kodlanmak zorundadır. Ayrıca ilk kez yapılan saldırıların tespit edilmesi mümkün değildir. Anormallik tespitinde ise önce normal erişimlerden normal kullanım profili bulunur ve profilden sapma gösterenler anormal olarak etiketlenir [5].

Kullanıcı davranış analizi ile nüfuz tespitinde, snort veri toplama işini, veri madenciliği ise veri analizini yapar. Veri madenciliği tekniklerinden biri olan kümeleme kullanılarak kullanıcılar için en uygun kümeler bulunur. Kendi kümesinden çok diğer kümelere benzerlik gösteren kullanıcıların sapma gösterdiğine, saldırgan olabileceğine karar verilir. Bu modelde kullanılan teknikler daha detaylı olarak aşağıda tanıtılacaktır.

2.1. SNORT Nüfuz Tespit Sistemi

Snort; kötüye kullanım tespiti modelinde çalışan bir nüfuz tespit sistemidir. Başlangıçta kötüye kullanım tespiti modelinde kural tabanlı bir nüfuz tespit sistemi olarak kullanıma sunulmuştur. Günümüzde trafik analizleri için takılabilir (plug-in) programcıklar kullanarak ağ verisi toplama, protokol başlıklarındaki anormalliklerin tespiti gibi işlemler için de kullanılmaktadır[1].

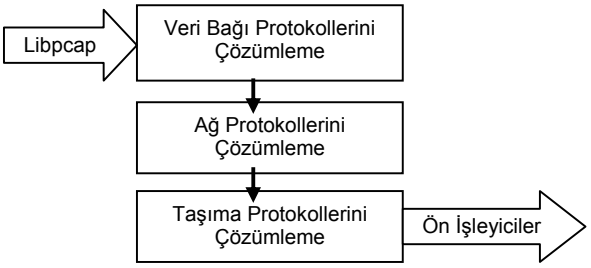
Snort çok katmanlı bir yapıdan oluşmaktadır. Belirli saldırıları tespit etmek ve istenen formatta çıktı vermek için bütün yapılarla bir arada çalışmaktadır. Snort tabanlı nüfuz tespit sistemlerinin mimarisi Şekil 1'de gösterilmiştir [1].



Şekil 1. Snort Nüfuz Tespit Sisteminin Mimarisi [1]

Paket Çözümleyicisi : Paket çözümleyicisi farklı tipteki ağ arabirimlerinden paketleri alır ve paketleri sonraki aşamalarda işlenmesi yada doğrudan tespit motoruna gönderilmek üzere hazır hale getirir. Snort; Ethernet, SLIP, FDDI, PPP gibi arabirimlerle birlikte çalışabilmektedir [1].

Ön İşleyiciler: Snort tarafından alınan bir paket, tespit motorunda gerçekleştirilecek kural uygulamaları öncesinde işlenmeye hazır hale getirilmelidir. Örneğin, paket parçalanmış bir yapıda ise paket boyutunun tespitinden önce tüm parçalanmış paketlerin yeniden bir araya getirilmesi gereklidir. Ön işleyicilerin görevi, tespit motorundaki farklı kuralların paketlere uygulanması için paketleri uygun bir hale getirmektir. Ek olarak, bazı ön işleyiciler anormalliklerin tespitinde ve veri paketlerindeki bilinen hataların tespitinde kullanılır [1].



Şekil 2. Çözümleyici Veri Akışı [1]

Tespit Motoru: Tespit motoru Snort'un en önemli parçasıdır. Tespit motorunun görevi; ağdaki bir paket yâda paket dizisiyle saldırı yapılıyor ise bunun tespit edilmesidir. Tespit için Snort kuralları kullanılmaktadır. Eğer bir paket herhangi bir kural ile eşleşirse, uygun eylem gerçekleştirilir aksi takdirde paket düşürülür. Uygun eylem, paketin kaydedilmesi yâda alarm verme olabilmektedir. [1]

Kayıt ve Alarm Verme Sistemi: Tespit motoru, ağdaki paketler içerisinde yapmış olduğu tespitlere bağlı olarak bu paketleri kaydedebilir yâda saldırı olarak öngördüğü paketleri alarm olarak çıktı modüllerine verebilir. Bu kayıtlar basit bir metin dosyasında, tcpdump formatında yâda diğer formatlarda saklanabilir. [1]

Çıkış Modülleri: Çıkış modülleri kaydetme ve alarm verme sistemleri tarafından üretilmiş çıktı tiplerinin kontrolünde kullanılır. [1]

2.2. Veri Madenciliği

Büyük miktardaki veri içerisinden anlamlı bilginin çıkarıldığı tekniğe veri madenciliği adı verilir. Veri madenciliği ile nüfuz tespiti birbirine yakın konulardır çünkü veri madenciliğinde yapılan işlerin bir kısmı anormal durumların tespiti ile ilgilidir [4].

Veri madenciliği kullanılarak geliştirilen nüfuz tespit sistemleri anormallik tabanlı yaklaşıma yakındır. Anormalliklerin tespiti için veri madenciliği yöntemleri kullanılabilir. Veri madenciliğinin doğasında yer alan özetlemeler yardımıyla veri indirgeme sayesinde gerçek zamanlı nüfuz tespitinin de yapılması mümkün olabilecektir. Ağ trafiğinden elde edilen normal trafik profilleri veri tabanlarında tutulur. Yeni girişimlerle normal trafik karşılaştırılarak nüfuz tespiti yapılır.

2.2.1. Kümeleme

Kümeleme benzerlik ve uzaklıklara dayalı olarak yerine getirilebilir. Benzerlik tabanlı kümelemede, küme içindeki elemanlar birbirine benzemektedir. Bu benzerlik küme dışındaki elemanlara benzerlikten daha fazladır. Uzaklık tabanlı kümelemede ise küme içindeki elemanlar birbirine daha yakın küme dışındaki elamanlara ise daha uzaktır.

Benzerlik ve uzaklık ölçütleri: Bir kümeye ait olan kayıt; o kümenin içindeki kayıtlara, dışındaki kayıtlara göre daha benzer veya daha yakındır. Veriler metrik olarak tanımlı ise, karakteristik değerler kullanılarak kümeler ifade edilebilir. Kullanıcı davranış analizi işlemi içinde kullanılan kümeleme k-NN algoritmasıdır. Burada amaç kullanıcının sapmasını kontrol etmektir.

3. Veri Madenciliği Tabanlı Kullanıcı Davranış Analizi ile Nüfuz Tespiti

Veri madenciliği tabanlı nüfuz tespitinin en önemli avantajı ağ trafiğindeki düzensizliklere bakarak daha önceden meydana gelmemiş saldırıları tespit edebilmesidir. Bunun için ağ üzerindeki veri trafiği eğitilen trafik bilgileriyle karşılaştırır ve düzensizlikler anormallik olarak tespit edilir.

KDA_NTM modeli, öncelikle ağ üzerindeki trafiği dinleyerek kullanıcıları özelliklerine (kullandığı port, veri miktarı, bağlantı saatleri, kullandığı protokoller vs.) göre kümelere atar. Veri madenciliği yaklaşımında, önce kullanıcıların normal davranışları öğrenilir ardından aktif kullanıcı erişimleri ile normal kullanım profili karşılaştırılarak tespit yerine getirilir. Bu yaklaşım kendi kendini eğitme tabanlı bir yaklaşımdır. Sistem çalıştığı sürece öğrenme işlemi de sürekli devam eder. Tespit işlemi için her zaman en güncel veri tabanı kullanılır.

Kullanıcılar tarafından alınan veya gönderilen veri paketleri analiz edilerek hangi kullanıcı anormal davranışlar gösteriyorsa o kullanıcı anormal kullanıcılar kümesine dâhil edilir. Bu çalışmada kullanıcı davranışları kullanıcıların sebep olduğu trafiğin istatistiksel özetleri ile sunulur. Analiz sırasında, kullanıcı davranışları saldırı niteliği kazandığı anda daha önceden belirlenen kısıtlamalar getirilir. Böylelikle daha önceden bilinmeyen saldırılar önlenmiş olur.

Veri madenciliği tabanlı kullanıcı davranış analizi şu aşamalardan oluşur.

- Bilgilerinin Toplanması
- Örüntülerinin Oluşturulması
- Analizlerinin Yapılması
- Kullanıcıları Kümelenmesi

Snort, modelin veri toplama aşamasında yer almakta olup modelin veri toplama işlemi Snort'un önişleme aracı ile kolay hale getirilecektir. Bilgi toplama aşamasında, hangi verilerin toplanacağına karar verilir

ve o bilgiler haricindeki veriler hesaba katılmaz. Bu işlem seçim olarak da bilinir. Seçim aşamasından sonraki aşama temizleme aşaması olup bu aşamada kullanım verisi ilişkisiz sahalardan temizlenir. Ardından kullanıcı tabanlı soyutlamalar yapılır. Kullanıcı tabanlı soyutlama için bir yöntem kullanıcı tabanlı verilerin özetlenmesidir. Örneğin, her bir kullanıcının iki saniye içerisinde almış olduğu veri paketi boyu bu bilgilerden birisi olabilir.

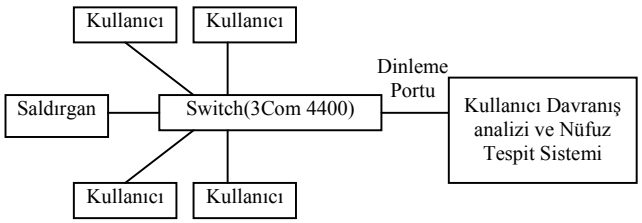
İkinci aşamada toplanan veriye istatistik, uyum kuralları, kümeleme, sınıflama ve sıralı örüntüler gibi teknikler uygulanarak örüntü oluşturma işlemi gerçekleştirilir.

Kullanıcı bilgileri daha önceden oluşturulan kullanıcı normal kullanım örüntüleri ile karşılaştırılarak analiz işlemi gerçekleştirilir.

Analiz işlemi sonrasında ağ içindeki kullanıcı bir sınıfa dâhil edilir. Daha sonra bu kullanıcıya atandığı sınıfının gerektirdiği güvenlik kuralları uygulanır.

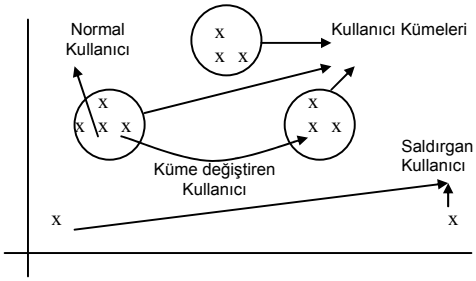
4. Gerçeklenen Model ve Çalışması

Bu modelde kullanıcı, MAC adresi ile tanımlı her bir ağ cihazıdır. Bir girişimin saldırı olup olmadığı kümeleme tabanlı anormallik tespiti ile bulunabilir. İlk kez meydana gelen saldırı normal erişimlerin bulunduğu kümeden farklılık gösterecek, farklı bir kümede yer alacaktır. Normal ile aradaki farklardan bilinmeyen girişimin saldırı kaynaklı olduğu tespit edilecektir. Bu işlemler yapılırken ağda merkezi anahtarlar üzerinden akan tüm trafik analiz edilecektir.



Şekil 3. Kullanıcı Davranış Analizi ve Nüfuz Tespit sistemleri

En yakın komşu algoritması ile kullanıcı davranışları kümelenir. Kimi kullanıcıların profilleri bağlı bulunduğu kümenin profiline yakın çıkacak kimileri ise kendi kümesinden farklı çıkacaktır. Bu modelde anormallikten kasıt bir kullanıcının kendi kümesi dışında farklı bir küme özellikleri göstermesidir. Bu da normal dışı uygulama olacaktır ve kullanıcının kendi kullanım yetkileri dışına çıktığı anlamına gelecektir. Bunlar içinde alarm verilip önlem alınması sağlanacaktır.



Şekil 4: Kullanıcı kümelerinin bulunması ve saldırganın tespiti

Modelde kullanılan veriler Snort tarafından elde edilmektedir. Snort'un sağladığı çıktılar kullanıcı bazlı olarak depolanır. Bu depolama işlemi yeterli miktarda eğitim verisi elde edene kadar devam eder. Elde edilen kullanıcı verileri kümeleme işleminden geçirilir. Kullanıcıların karakteristik özelliklerine bakarak kullanıcı tanımlaması yapılır. Bu özellikler gönderilen ve alınan paket yapıları, kullanılan protokoller, veri büyüklükleri, erişim zamanları, kullanılan portlar gibidir. Kullanıcı karakteristiği belirleyen özellikleri ne kadar çoğaltırsak modelin güvenirliği o derecede artacaktır. Bunun dezavantajı ise işlem yükü artacağından kullanılan PC'nin hızlı işlem gücüne sahip olması gerektirmesidir. Benzer kullanıcılar bir kümede toplanır. Kümeleme sırasında farklı veriler, kümeleme dışında tutularak ideal kümeler oluşturulur. Daha sonra bu kullanıcı verileri ait olduğu küme normal alınarak kontrol edilir. Anormal davranışlar gözlendiği anda uyarı sistemi devreye girer.

Sistem eğitildikten sonra eş zamanlı kullanıcı analizleri yapılır. Eğer kullanıcı trafiği belirlenen esasları sağlamıyorsa şüpheli konumuna düşürülür ve hareketleri detaylı incelemeye alınır. Belirli süre incelendikten sonra hala aynı konum devam ediyorsa kullanıcı saldırgan konumuna alınır ve bu kullanıcıdan gelen paketler engellenir. Kullanıcı davranışları anormallik gösterdiği anda nüfuz tespit sistemi devreye girecektir.

5. Sonuç ve Öneriler

Bilgisayar sistemlerinin tam olarak korunmasını sağlamak amacıyla sistem gerçek zamanlı olarak çalışmaktadır. Bir bilgisayar sistemini korumanın en iyi yolu çeşitli güvenlik araçlarını birlikte kullanmaktır.

Bu çalışmada önerilen model, bilinen Nüfuz tespit sistemlerine veri madenciliği tekniklerinin uygulanmasını ve gerçek zamanlı kullanıcı analizinin yapılmasını önermektedir. Böylece hem nüfuz işlemi gerçekleşmeden saldırgan etkisiz hale getirilir. Hem de ağda bulunan yetkisiz kullanıcılar engellenmiş olur. Bu modelin getireceği en önemli yenilik daha önceden

bilinmeyen saldırıları da anormallik tespiti yaklaşımıyla tespit etmesidir.

Kullanıcı davranışlarını belirlerken, kullanıcının veri akışı, kullandığı portlar, aldığı veri paketleri, kullanım zaman aralıkları, paket büyüklükleri ve kullandığı sistem çağrıları değerlendirilir. Tüm bu veriler toplandıktan sonra nüfuz tespit sistemi içerisinde kullanılır. Bu işlem virüs, trojan ve ağda oluşturulacak saldırı niteliği taşıyan diğer programları da engelleyecektir.

Modelin geliştirilmesi sayesinde dağıtık yapıda çalışma mümkün hale getirilerek modelin etki alanı genişletilebilir. istemci ve sunumcu arasında belli bir iletişim kurularak merkezi bir veri tabanı oluşturulabilir. Bu iletişimin kötü niyetli kişiler tarafından dinlenmesi mümkün olabilir. Bu amaçla istemci ve sunumcu arasındaki iletişim bazı şifreleme algoritmaları ile şifrelenebilir.

Kaynakça

[1] J. Koziol, (2003), "Intrusion Detection with Snort", Indiana, USA, Sams Publishing, 23-68.

[2] Wenbao Jiang, H.Song, Y.Dai, "Real-Time Intrusion Detection for High-Speed Networks" Computers and Security 2004

[3] J.M. Stanton, K.R. Stama, P. Mastrangelob,J.Joltonb, "Analysis of end user behaviours" Computers and Security 2004

[4] Sunita Sarawagi, "Intrusion Detection Using Data Mining Techniques", 2001

[5] H.Takci, İ.Soğukpınar, "Saldırı Tespitinde Yeni Bir Yaklaşım",19.TBD Bilişim Kurultayı, 3-6 Eylül 2002 ,İstanbul

[6] Turker Akyüz, İbrahim Soğukpınar, "AKAT Akıllı Açıklık Tarayıcıları", Türkiye Bilişim Derneği, 21. Ulusal Bilişim Kurultayı, 2004 ODTÜ-ANKARA

[7] Qin Degang, Zeng Zhongtao, Margarita C. S. Paterno, Web Usage Mining

[8] Yongjian Fu Kanwalpreet, Sandhu Ming-Yi Shih, 1999, Clustering of Web Users Based on Access Patterns

[9] Mustafa Coşkun, İbrahim Soğukpınar, "Dağıtık Saldırı Tespit Sistemleri için bir Model", 19. Bilişim Kurultayı, İstanbul-Turkey, 2002.

[10] Oren Etzioni, 1996, Quagmire or Gold Mine, Communications of the ACM