**T.C.**

**SAKARYA ÜNİVERSİTESİ**

**BİLGİSAYAR VE BİLİŞİM BİLİMLERİ FAKÜLTESİ**

BSM 401 BİLGİSAYAR MÜHENDİSLİĞİ TASARIMI

SALDIRI TESPİT SİSTEMLERİNDE  
VERİ MADENCİLİĞİ

G191210386 Yasin ALTUNBAŞAK

|  |  |  |
| --- | --- | --- |
| Bölüm  Danışman | :  : | BİLGİSAYAR MÜHENDİSLİĞİ  Doç. Dr. Nilüfer Yurtay |

2021-2022 Yaz Dönemi

**ÖNSÖZ**

Geçmişten günümüze bilgisayar ve teknoloji alanında büyük gelişmeler gerçekleşmiş ve gerçekleşmeye devam etmektedir. Bu gelişmelerle birlikte verilerin ve sistemlerin güvenliğini tehdit eden birçok saldırıyı da beraberinde gelmektedir. Sistemler ve verilerim güvenliği için de saldırıların nasıl ve ne zaman gerçekleştiğini kavramak Siber Güvenliğin en önemli gerekliliklerinden birisidir.

Saldırı Tespit sistemlerinin kullandığı yöntemlerde Veri Madenciliği ön plana çıkmaktadır. Özellikle SIEM uygulamalarında saldırılar sonucunda tutulan büyük boyutta verilerin ihtiyaca uygun bir şekilde ayıklanması, Veri Madenciliğinin gereksinimlerindendir.



Şekil 1.1. Sakarya Üniversitesi logosu

**ÖZET**

Anahtar kelimeler: Veri Madenciliği, Siber Güvenlik, Saldırı Tespit Sistemleri

Siber Güvenlik alanının başlıca

Saldırı tespiti için birçok teknik kullanılmaktadır, bu tekniklerden birisi de büyük miktardaki veriden anlamlı bilgiyi bulmaya yarayan veri madenciliği tekniğidir. Sunucu günlüklerindeki erişim bilgilerine dayalı olarak kullanıcı ve onun aktivitelerini tanımayı amaçlayan bu teknik saldırı tespitinde etkin olarak kullanılabilmektedir.

Bu tasarım çalışması, saldırı tespit sistemlerinde veri madenciliğinin önemini ve detaylarını inceleyerek, SIEM sistemlerini kavramamıza yardımcı olacaktır. İnceleyeceğimiz çalışma WEKA aracı üzerinde gerçekleştirilmiş bir Web Log önizleme süreci ve analizidir.

# GİRİŞ

Saldırı Tespit Sistemleri, ağlara veya sistemlere karşı yapılan kötü niyetli aktiviteleri ya da sistemin politika ihlallerini izlemeye yarayan cihaz ya da yazılımlardır. Tespit edilen herhangi bir aktivite veya ihlali, ya bir yöneticiye bildirilir ya da bir güvenlik bilgi ve olay yönetimi (SIEM) sistemi kullanılarak merkezi olarak toplanır. SIEM sistemi, çeşitli kaynaklardan gelen çıktıları birleştirir ve kötü niyetli alarmı yanlış alarmlardan ayırmak için alarm filtreleme teknikleri kullanır.

Veri Madenciliği ise bu sistemlerin iç yüzünde büyük rol oynamaktadır. Bu sistemin birlikte çalışmasını incelemeden önce Saldırı Tespit Sistemleri ile Veri Madenciliğini ayrı dallarda tanımamız gerekir.

* 1. Saldırı Tespit Sistemleri

Olası atak, ihlal ve tehditlerin tespiti için ağda gerçekleşen aktivitelerin izlenmesi ve trafiğin analiz edilmesini sağlayan yöntemleri kapsayan saldırı tespit sistemleri önemli güvenlik bileşenlerinden biridir. Saldırı engelleme sistemleri ise saldırıların saptanıp önlenmesi işlemlerini kapsayan ağ güvenlik sistemleridir.

Günümüzde ağların kompleks bir yapıya sahip olması, internet başta olmak üzere diğer ağlara birçok erişim noktası ile bağlı olması, siber saldırıların her geçen gün çeşitlenmesi ve artması, aynı zamanda bu karmaşık ağ sistemlerinin artık sadece şifreleme veya güvenlik duvarı ile korunamayacağı gibi gerçeklikler ağ trafiğinin devamlı izlenip saldırı girişimlerinin gerçek zamanlı olarak tespitini kaçınılmaz kılmıştır.

IDS/IPS sistemleri ağı sık sık monitör etmek, olası tehditleri tanımlamak ve bunlarla ilgili olay kayıtlarını(logları) tutmak, saldırıları durdurmak ve güvenlik yöneticilerine raporlamak gibi işlevlere sahiptir. Bu sistemler bazı durumlarda kurumların güvenlik politikalarındaki zayıflıkları ortaya çıkarmak için de kullanılabilmektedir. IDS/IPS aynı zamanda saldırganların ağla ilgili bilgi toplama faaliyetlerini algılayarak saldırganları bu erken aşamada da durdurabilme işlemini gerçekleştirebilirler.

1.1.2 IDS/IPS

Özellikle de son yıllarda artan gelişmiş siber saldırılara karşı yeni nesil Firewall cihazları ile bütünleşik olarak kullanılan IPS ve IDS sistemleri ciddi zararlara yol açan saldırıların tespiti, yakalanması ve önlenmesi konularında doğru konfigüre edildiğinde önemli bir rol oynamaktadır. IDS derinlemesine paket analizinde saldırıyı tespit ederek loglama işlevi yürütürken, IPS sistemleri saldırıları öğrenerek veya kural listesi üzerinden tespit ederek engellemektedir.

1.1.1.2 IDS

IDS (Intrusion Detection System) saldırı tespit sistemi olarak adlandırılan temelde ve bir ağı izleyip güvenlik açıkları ve saldırıları tespit etmek üzere kurgulanmış yazılım veya donanım şeklinde çalışan siber güvenlik sistemidir. IPS sistemi ile kombine bir şekilde veya ayrı olacak şekilde kurgulanabilir. Her iki sistemde güvenlik duvarının arkasında kurgulanmaktadır. IDS sistemleri saldırıları imza alarmı denilen yöntemle belirlemektedir. Signature (İmza) denilen şey ise saldırıları tespit etmek amacıyla önceden belirlenmiş durumlarla karşılaşıldığında tetiklenen birtakım kurallar dizisidir. Kurgulanan her signature farklı şekilde davranabilir.

IDS sistemlerini bilgisayar ve ağlara yönelik oluşturulabilen iki sensör çeşidi vardır. Network sensörü ağ üzerinde kurgulanarak ağdaki kötü amaçlı bağlantılar ve veri paketlerini tespit edip bloke eder. Server sensör ise sunucu üzerinde konumlandırılır ve sunucunun veri trafiğini izler. Sunucu üzerindeki yetkisiz ve zararlı işlemleri tespit ederler. Her iki sensör tipi de aynı veritabanına kullanırlar. Bunun yanında IDS sistemleri algılama çeşidine göre de sınıflandırılabilir.

(a) Anormal temelli IDS

Bir ağ üzerindeki normal çalışma sisteminin dışında gerçekleşen anormal durumları izlemek üzerine kurulmuştur. Önceden belirlenmiş bir davranış modeli olmasa bile sistemde normal olarak değerlendirilmiş hareketler haricindeki olayları anormal diye nitelendirir. Yanlış negatif alarm verme durumu oldukça fazla olsa da tanımlanmamış saldırıları yakalamanın önemli bir yoludur.

(b) İmza temelli IDS

İmza bazlı yani daha önceden bilinen ve belirtilmiş kurallar ve davranışlar çerçevesinde çalışan saldırı önleme sistemidir. Olumsuz yanı ise daha önceden belirlenmemiş saldırı modellerini yakalayamamasıdır.

**1.1.1.3. IPS**

İzinsiz Giriş Önleme ya da Saldırı Önleme Sistemi (IPS), bir ağı kötü amaçlı etkinliklere ve saldırılara karşı sürekli olarak izleyen ve bu saldırı gerçekleştiğinde raporlama, engelleme veya devre dışı bırakma işlemleri dahil olmak üzere, bu saldırıyı önlemek için harekete geçen bir ağ güvenlik aracıdır (bu, bir donanım aygıtı veya yazılım olabilir). İzinsiz giriş sistemleri, birçok ağ güvenliği teknolojisi gibi ağ performansını düşürmeden yüksek hacimli trafiği tarayabilecek kadar güçlü olmalıdırlar.

IPS, yönlendirilen ağ trafiğini kötü niyetli faaliyetler ve bilinen saldırı kalıpları için aktif olarak tarayarak çalışır. IPS motoru ağ trafiğini analiz eder ve bilinen saldırı kalıpları için bit akışını dahili imza veri tabanıyla sürekli olarak karşılaştırır. Bir IPS, kötü niyetli olduğu belirlenen bir paketi bırakabilir ve bu eylemi, saldırıyı yapan kişinin IP adresinden veya bağlantı noktasından gelecek tüm trafiği engelleyerek takip edebilir. Kötü niyetli olmayan normal trafik ise herhangi bir kesinti olmaksızın devam edebilir.

Saldırı Önleme Sistemleri, trafik kalıplarını veya paketlerini izlemek ve bunlara tepki vermek gibi daha karmaşık gözlem ve analizler de yapabilir. Algılama mekanizmaları şunları içerebilir:

* Adres değiştirme
* Genel desen eşleştirme
* TCP bağlantı analizi
* Paket anormalliği algılama
* Trafik anormalliği algılama
* TCP/UDP bağlantı noktası eşleşmesi

Bir IPS tipik olarak gözlemlediği olaylarla ilgili bilgileri kaydeder, güvenlik yöneticilerini bilgilendirir ve raporlar üretir. Bir ağın güvenliğini sağlamaya yardımcı olmak için bir IPS, ortaya çıkan internet tehditlerini sürekli olarak izlemek ve engellemek için önleme ve güvenlik güncellemelerini otomatik olarak alabilir.

1.1.1.4 IPS ile IDS arasındaki farklar

IPS bir kontrol sistemidir kendi veritabanı üzerinde belirlenen kural setlerine göre oluşan saldırıları tespit ettiğinde veri paketlerini reddederek veri alışverişini durdurur. Veritabanın sürekli güncel tutulmasına ihtiyacı vardır. Her iki sistemde bilinen tehditleri kendi ve veritabanları üzerinde karşılaştırarak bulmaya çalışır. IDS bir saldırı tespit ve monitörize etme sistemidir. Kendi başlarına aksiyon almazlar. Aksiyon almak için başka bir sistem veya bir personele ihtiyaç duyar.

|  |  |  |
| --- | --- | --- |
| **Karşılaştırma Parametreleri** | **IDS** | **IPS** |
| **Full – Form** | Saldırı Tespit Sistemleri | Saldırı Önleme Sistemleri |
| **Çıkış Tarihi** | IDS, 1984 ile 1986 arasında piyasaya sürüldü. | IPS 2000’lerin ortalarında piyasaya sürüldü. |
| **Tanım** | IDS, herhangi bir kötü amaçlı yazılım için ağ geçidinden geçen dosyaları algılayan sistemdir. | IPS, sağlanan kural kümesine göre algılanan kötü amaçlı yazılımları algılayan ve hatta çözen yazılımdır. |
| **Tür** | IDS, pasif bir yazılım türüdür. | IPS aktif bir yazılımdır. |
| **Çalışma** | IDS, kötü amaçlı yazılım ve hataların algılanmasını ve bildirilmesini içerir. | IPS, sorunu kendi başına çözdüğü için insanların veya diğer yazılımların katılımını gerektirmez. |
| **Performans** | IDS, çevrimiçi olmayan dağıtım nedeniyle ağ performansını etkilemez. | IPS, algılama işlemi nedeniyle ağı yavaşlatır. |
| **İletişim** | IDS’nin iletişimi bant dışındadır. | IPS veri iletişimi ile uyumludur. |
| **Avantaj** | IDS, kötü niyetli trafiği algılar. | IPS kötü niyetli trafiği düşürür, uyarır veya temizler. |

Tablo 1.1. IPS IDS karşılaştırması

**1.2. Veri Madenciliği**

Veri Madenciliği, büyük miktarda veri setleri içerisinden birbirleri ile anlamlı bilgileri alıp gelecek hakkında en doğru tahminde bulunulmasını ve birçok keşfedilmemiş bilgiyi açığa çıkarmada kullanılır.

Veri madenciliğinde birçok yöntem kullanılmıştır ve bunlar problem türüne, uygulama alanlarına ve eldeki veriye göre çeşitlilik göstermektedir. Veri madenciliğinin en önemli amacı ise bir araya getirilmiş verilerin kendi içlerinde istatistiksel yöntemler ile incelenip istenilen veya ilgili kuruma kullanılması için değerlendirilip, kullanıma başlanmasıdır.

**1.2.1. Veri Madenciliğine Giriş**

Veri madenciliği, büyük boyutlu veri ambarlarının oluşmasının bir sonucudur.

Günümüzde kurumlar büyük miktarlarda veri üretmekte ancak bu veriler içerisinde

anlamlı ve yararlı bilgiyi ortaya çıkarmakta zorluklar yaşamaktadırlar. Geleneksel

istatistiksel yöntemlerle büyük boyuttaki veriyi çözümlemek kolay değildir. Bu

nedenle verileri işlemek ve çözümlemek için özel yöntemlere gereksinim

duyulmuştur. Veri madenciliği yöntemleri bu gereksinimi karşılamak üzere ortaya

çıkmıştır.

Bilgisayar teknolojilerinin sağlamış olduğu çok hızlı veri işleme ve yüksek hacimde

veri depolama imkânları yardımıyla ve farklı disiplinlerin katkısıyla sağlanan

araçlarla, sahip olunan çok büyük hacimlerdeki veriden, karar vericinin etkin ve daha

fazla bilgiye dayalı karar vermesinde kullanabilmesi için önceden bilinmeyen, gizli,

örtük, klasik metotlarla ortaya çıkarılması güç, faydalı, ilginç, anlaşılabilir; ilişki,

örüntü ve bağıntıların otomatik veya yarı otomatik bir şekilde ortaya çıkarılması

olarak tanımlanır.

**1.2.2. Kavram**

Veri madenciliği deyimi yanlış kullanılan bir kavram olabileceğinden buna eş değer başka kullanımlar da literatüre geçmiştir. Veritabanlarında bilgi madenciliği, bilgi çıkarımı veri ve örüntü analizi, veri arkeolojisi gibi.[3] Bu terimler arasında "Veritabanlarında Bilgi Keşfi" en yaygınıdır.[4] Alternatif olarak veri madenciliği aslında bilgi keşfi sürecinin bir parçası şeklinde kabul görmektedir. Bu adımlar:

* Veri temizleme (gürültülü ve tutarsız verileri çıkarmak)
* Veri bütünleştirme (birçok veri kaynağını birleştirebilmek)
* Veri seçme (yapılacak olan analizle ilgili olan verileri belirlemek )
* Veri dönüşümü (verinin veri madenciliği tekniğinden kullanılabilecek hale dönüşümünü gerçekleştirmek)
* Veri madenciliği (veri örüntülerini yakalayabilmek için akıllı metotları uygulamak)
* Örüntü değerlendirme (bâzı ölçümlere göre elde edilmiş bilgiyi temsil eden ilginç örüntüleri tanımlamak)
* Bilgi sunumu (mâdenciliği yapılmış olan elde edilmiş bilginin kullanıcıya sunumunu gerçekleştirmek).[5]

**1.2.3. Yöntem**

Veri madenciliği adımı, kullanıcı ve bilgi tabanıyla etkileşim halindedir. İlginç örüntüler kullanıcıya gösterilir, ve bunun ötesinde istenirse bilgi tabanına da kaydedilebilir. Buna göre, veri madenciliği işlemi, gizli kalmış örüntüler bulunana kadar devam eder.

Bir veri madenciliği sistemi, aşağıdaki temel bileşenlere sahiptir:[6]

* Veritabanı, veri ambarı ve diğer depolama teknikleri
* Veritabanı ya da Veri Ambarı Sunucusu
* Bilgi Tabanı
* Veri Madenciliği Motoru
* Örüntü Değerlendirme
* Kullanıcı Arayüzü

Veri madenciliği, eldeki verilerden üstü kapalı, çok net olmayan, önceden bilinmeyen ancak potansiyel olarak kullanışlı bilginin çıkarılmasıdır. Bu da; kümeleme, veri özetleme, değişikliklerin analizi, sapmaların tespiti gibi belirli sayıda teknik yaklaşımları içerir.

Başka bir deyişle, veri madenciliği, verilerin içerisindeki desenlerin, ilişkilerin, değişimlerin, düzensizliklerin, kuralların ve istatistiksel olarak önemli olan yapıların yarı otomatik olarak keşfedilmesidir.

Temel olarak veri madenciliği, veri setleri arasındaki desenlerin ya da düzenin, verinin analizi ve yazılım tekniklerinin kullanılmasıyla ilgilidir. Veriler arasındaki ilişkiyi, kuralları ve özellikleri belirlemekten bilgisayar sorumludur. Amaç, daha önceden fark edilmemiş veri desenlerini tespit edebilmektir.

Veri madenciliğini istatistiksel bir yöntemler serisi olarak görmek mümkün olabilir. Ancak veri madenciliği, geleneksel istatistikten birkaç yönde farklılık gösterir. Veri madenciliğinde amaç, kolaylıkla mantıksal kurallara ya da görsel sunumlara çevrilebilecek nitel modellerin çıkarılmasıdır. Bu bağlamda, veri madenciliği insan merkezlidir ve bazen insan – bilgisayar arayüzü birleştirilir.

Veri madenciliği sahası, istatistik, makine bilgisi, veritabanları ve yüksek performanslı işlem gibi temelleri de içerir.

**1.3. Saldırı Tespit Sistemlerinde Veri Madenciliği**

Güvenlik Duvarları ve buna bağlı SIEM araçları sistemde gerçekleşen hareketleri tutmakta, yetkili IT ekibi tarafından yazılan korelasyon kuralları sonucu alarm oluşturabilme özelliğine sahiptir. Bu tutulan bilgiler “log” adı ile bilinir.

Log dosyaları bilgisayar sisteminin tarih kitabı olarak kabul edilmektedir. Loglar sistemin çalışmasıyla ilgili önemli bilgiler içermektedir. Bu bilgiler genellikle performans ölçümü, işlemsel profil tespiti, anormalliklerin tespiti, hata ayıklama, güvenlik tehditlerinin tespiti gibi işlemlerde kullanılmaktadır. Son zamanlarda çeşitli veri madenciliği ve makine öğrenme algoritmaları log dosyalarındaki bilgileri analiz etmek üzere yoğun olarak kullanılmaktadır.

Log dosyaları çok fazla bilgi içermektedir, ancak kullanılan dosya formatından bu bilgilerden çıkarım yapmak oldukça zordur. Dolayısıyla verileri analiz etmeye yarayacak bir araca gereksinim duyulmaktadır [9]. Önceleri log analizleri sistem yöneticileri tarafından manuel olarak gerçekleştirilmiştir. Bu durum bazı olayların gözden kaçmasına sebep olabileceği gibi; her geçen gün artan veri miktarı manuel yöntemlerle analizin yetersiz kalmasına sebep olmuş ve analistler çeşitli araçlar

geliştirmeye başlamışlardır [10]. Son yıllarda log yönetim sistemleri kurumlar tarafından yaygın olarak kullanılmaya başlanmıştır. Bazı şirketler (IBM, MacAfee ve Splunk gibi) kendilerine özel log yönetim çözümlerini sunmaktadır. Ancak bu sistemlerin uygun donanım gerektirmesi ve log verilerini analiz etmek üzere web kullanım madenciliğini içermemesi problem olmaktadır.

log dosyaları serverların kullanımıyla ilgiliverileri depolamaktadır. Bu verilerin analizi son derece önemlidir. Genellikle log dosyaları farklı formatlarda kaydedilmekte bu durum analiz edilmelerini zorlaştırmaktadır.

log dosyalarında 2 farklı format görülmektedir.Bunlardan biri “Common Log Format” ve diğeri ise “Combined Log Format”tır. Bu makalede çalışılan erişim dosyasındaki kayıt formatı ikinci türdendir. Aşağıda bu formata uygun log erişim dosyasının alanları sıralanmıştır.

* İsteği gerçekleştiren IP numarası,
* İsteğin yapıldığı tarih ve saat,
* İstek yapılan URL,
* Sunucunun verdiği yanıt olan durum kodu,
* Gönderilen dosyasının boyutu,
* İsteğe nereden ulaşıldığı,
* İstekte bulunan kullanıcıya ait tarayıcı bilgisi

# SİSTEMATİK YAKLAŞIM

Bu kısımda Bölüm birde bahsettiğimiz WEKA açık kaynak veri madenciliği aracını, deneysel çalışma bölümünü anlamakta kolaylaştırma amacıyla tanıtacağım bölümdür.

* 1. Weka

Weka Waikato Üniversitesi tarafından geliştirilmiş açık kaynak bir veri madenciliği aracıdır. Weka kullanıcıları sunduğu grafik ara yüz sayesinde tüm işlemlerin çok daha kolay bir şekilde yapılmasını sağlamaktadır. Weka üzerinde veri setleri üzerinde işlem gerçekleştirmek için birçok paket barındırmaktadır. Veri setleri üzerinde önişlem yapılması, sınıflandırma ve kümeleme işlemleri için, birliktelik kurallarının tanımlanması, özellik çıkarımı ve görselleştirme işlemleri için ayrı paketler bulunmaktadır. Veri madenciliği uygulamalarında kullanılan daha birçok uygulama programı bulunmaktadır. Rapid Miner ve R bunlardan en yaygın olarak kullanılan programlar arasındadır. Bu programların birbirine göre avantaj ve dezavantajları bulunmaktadır.

WEKA verileri basit bir dosyadan okur ve veriler üzerindeki stokastik değişkenlerin sayısal veya nominal değerler olduğunu kabul eder. Aynı zamanda veritabanı (database) üzerinden de veri çekebilir ancak bu durumda verilerin bir dosya verisi şeklinde olması beklenir.

WEKA üzerinde makine öğrenmesi ve istatistik ile ilgili pekçok kütüphane hazır olarak gelmektedir. Örneğin veri ön işlemesi (data preprocessing), ilkelleme (regression), sınıflandırma (classification), gruplandırma (clustering), özellik seçimi veya özellik çıkarımı (feature extraction) bunlardan bazılarıdır. Ayrıca bu işlemler sonucunda çıkan neticelerinde görsel olarak gösterilmesini sağlayan görüntüleme (visualization) araçları bulunmaktadır.

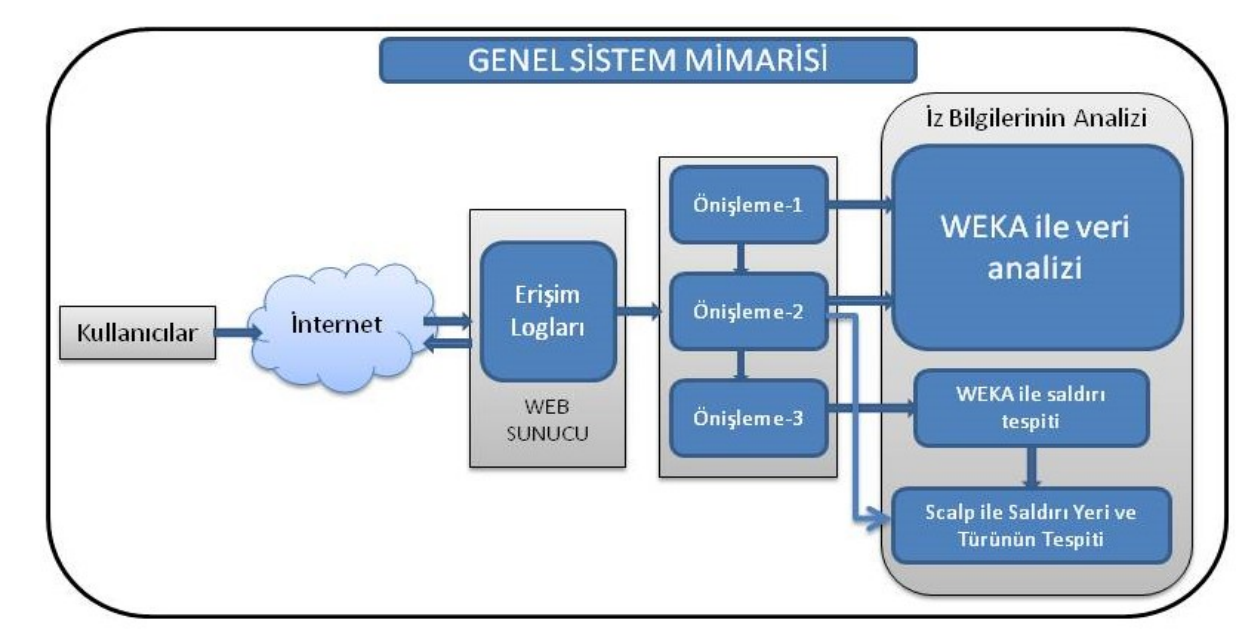
# DENEYSEL ÇALIŞMA

WEKA aracını hali hazırda gerçekleştirilmiş bir Web sayfası üzerinde kullanımı bu bölümde inceleyeceğiz.

* 1. Deney Düzeneği

Çalışmada bir yemek web sitesine ait web erişim log dosyası 3 farklı yöntemle analiz edilmiştir.

1. WEKA yazılımı kullanılarak web erişim loğlarından istatistiksel analizler yapılmıştır.
2. WEKA yazılımı kullanılarak saldırı tespitine yönelik çeşitli çıkarımlarda bulunulmuştur.
3. WEKA’dan elde edilen sonuçlar ışığında filtrelenmiş veri kullanılarak web saldırı tespitinde kullanılan Apache Scalp aracında saldırı ve açıklık tespiti gerçekleştirilmiştir.

Çalışmanın genel sistem mimarisi Şekil 4.1.’de sunulmuştur:

Şekil 4.1. WEKA genel sistem Mimarisi

* 1. Önişleme Adımları

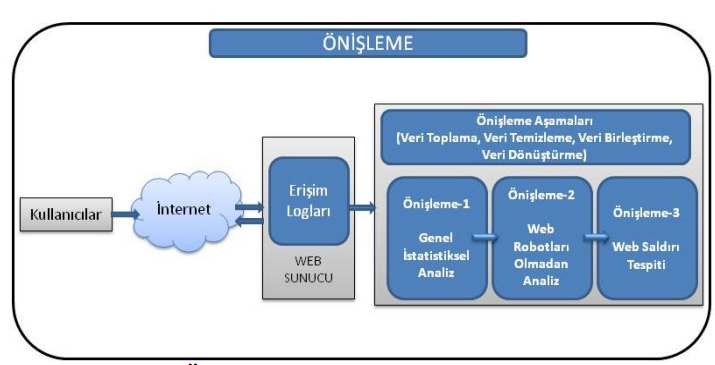
Veri analizinde yapılacak çalışmanın kalitesi bu aşamayla doğrudan ilişkilidir [34]. Bu aşama web kullanım madenciliğinin en önemli aşaması olarak kabul edilmektedir; çünkü etkili bir şekilde yapıldığında zaman ve kaynak tasarrufu sağlamaktadır [35].

Veri önişleme web madenciliğinde amaca yönelik olarak değişkenlik göstermektedir. Log dosyasında stil bilgileri ve grafik bilgileri gibi önemli olmayan nesneler atılarak veri temizlemesi yapılmıştır [36]. Ancak eğer resim ağırlıklı siteler ve haber siteleri söz konusu olursa bu bilgilerin atılması doğru olmayabilir [37].

Web kullanım madenciliğinde özellikle 200 serili durum kodlarının yer aldığı isteklerin analiz edilmesi gerekmektedir. Web loglarından saldırı analizi yapılacağında ise web kullanım madenciliği önişlemesinin aksine anomali tespiti içerebileceğinden 400 serili durum kodlarının korunması gerekmektedir. 200 serili isteğin başarılı olduğunu gösteren durum kodlarının yer aldığı isteklerin ise silinmesi gerekmektedir. Ancak saldırı başarılı olduğunda hata kodlarının ardından başarı durum kodu alınacağı dikkate alınmalıdır [25].

Önişlemenin etkin yapılması analizin hızını, güvenilirliğini, genel olarak performansını ciddi oranda etkilemektedir. Veri madenciliği yöntemleri ile web saldırı tespiti yapan çalışmaların [38, 39, 40] birçoğunda önişleme aşamasında hata durum kodları veya başarı durum kodlarına yönelik önişleme çalışması yapılmadığı görülmüştür.

Bu çalışmada web madenciliği amacına yönelik olarak önişleme 3 aşamada gerçekleştirilmiş; her bir aşamaya özel önişleme dosyası oluşturularak web madenciliği analizi yapılmıştır. Şekil 4.2.’de bu aşamalar gösterilmektedir.



Şekil 4.2. Önişleme aşamaları

Her bir aşamaya özel gerçekleştirilen önişleme bilgileri aşağıda sunulmuştur. 1. Web kullanım madenciliğine yönelik genel istatistiksel sonuçlar çıkarma amacıyla gerçekleştirilen önişleme adımları aşağıda sunulmuştur:

* Kimlik ve kullanıcı alanlarında veri tutulmadığı için bu alanlar silinmiştir.
* Resim ve müzik dosyaları silinmiştir.
* WEKA’da yapılacak örüntü keşfi aşamaları için tarih formatı “gün/ay/yıl” formatına dönüştürülmüştür.
* Metot, protokol ve URL alanları birleştirilmiştir.

2. İlk maddede gerçekleştirilen önişlemede web robotları erişimleri de bulunmaktadır. Web robotları (crawler, spider, bot vb.) “robots.txt” dosyasına erişmekte ve yöneticinin verdiği izinlere göre işlem yapmaktadırlar. İz bilgilerinin kaydadeğer bir bölümünü web robotları oluşturduğundan bu isteklerin temizlenmesi analizde kullanıcı odaklı sonuç çıkarma açısından çok önemlidir.

3. Web saldırı tespitine yönelik analiz yapmak amacıyla ilk iki maddede gerçekleştirilen önişlemeye ilave olarak 400 serili hata durum kodlarının dışındaki durum kodlarının yer aldığı istekler silinmiştir.

**4.3.1 Weka ile Örüntü Keşfi ve Analizi**

(Pattern Discovery and Analysis by using Weka) WEKA makine öğrenme ve veri madenciliği için kapsamlı bir araçtır [41]. WEKA, açık kaynak kodlu ve java platformu üzerinde geliştirilmiş bir programdır [42]. WEKA’ya verilerin yüklenmesi için çeşitli yöntemler bulunmaktadır:

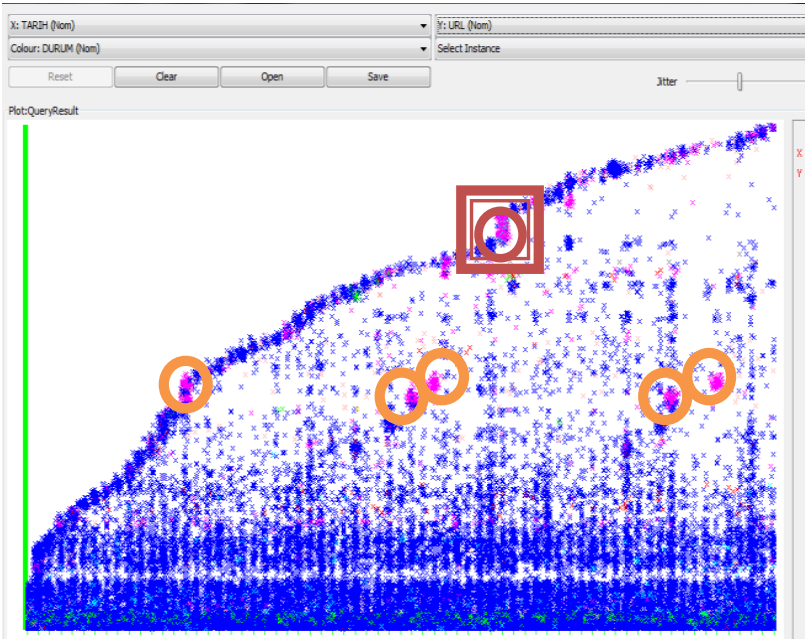
* Verileri dosyadan açma
* Verileri URL’den açma
* Verileri veritabanından açma

Bu çalışmada veriler veritabanı aracılığı ile WEKA’ya aktarılmıştır. 4.2. Önişleme-1 verisi ile analiz (Analysis with preprocessing-1 data) WEKA ile örüntü keşfi ve analiz için önişleme-1 adımından elde edilen veri tablosu, veritabanı bağlantısıyla yüklenmiştir. WEKA’ya veriler yüklendikten sonra WEKA içerisinde yer alan önişlem adımında görüntülenmektedir. Bu aşamada filtreleme işlemleri ve istatistiksel analizler yapılabilmektedir. Erişim log dosyasının analizi sonucunda ortaya çıkan istatistiksel rakamlar Tablo 4.3’de listelenmiştir.

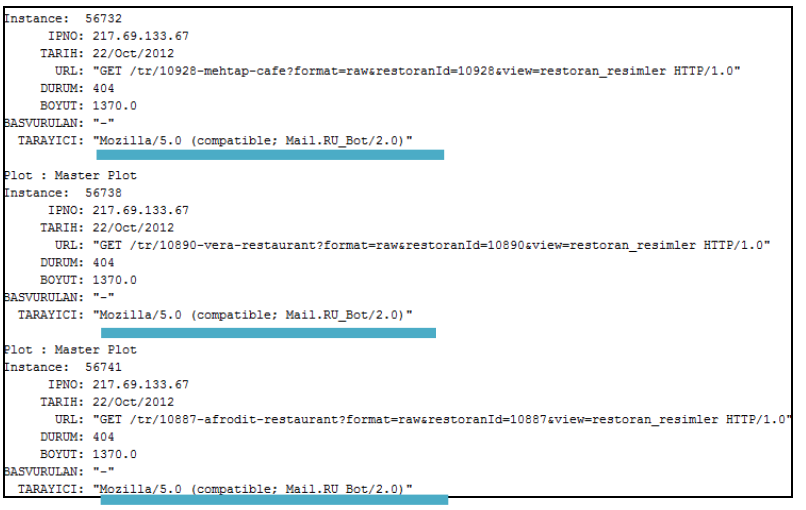


Tablo 4.1. Erişim log dosyası istatistiksel sonuçlar

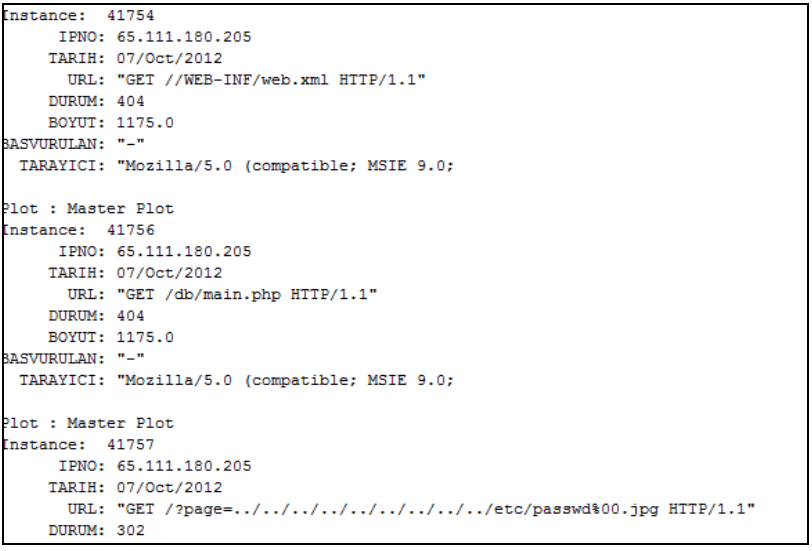
WEKA’nın Visualize sekmesinden analiz edilen verinin istenilen alanları arasındaki ilişki yayılım grafiği görülebilmektedir. Bu sekmede alanların dağılımlarına göre kesişim noktaları görülebildiğinden kümeleme analizine olanak sağlamaktadır . İncelenen ana log verisi için Tarih, URL ve durum alanlarına göre işlem yapıldığında Şekil 4.3’deki sonuç elde edilmiştir. Burada X ekseni tarih alanını, Y ekseni URL alanını, ekranın en altında görüldüğü gibi farklı renkler de durum alanını temsil etmektedir. Bu 3 alana göre pembe renk ile temsil edilen 404 hata durum kodu için Şekil 4.4'de işaretlenen 6 yerde küme oluştuğu görülmektedir.

 Şekil 4.3. WEKA’nın Visualize sekmesindeki analiz

Visualize ekranında koordinat düzleminde yer alan her bir kaydın üstüne tıklandığında detaylı bilgi görüntülenmektedir. Bu kümelerin bulunduğu alanlar bu şekilde detaylı incelendiğinde çoğu kümenin arama robotlarının kısa sürede arka arkaya gerçekleştirdiği indeks kayıtlarından oluştuğu görülmektedir. Şekil 4.5.’de tarayıcı alanları incelendiğinde isteklerin web robotlarından geldiği görülmektedir.

 Şekil 4.4. Analiz ile ilgili detaylı bilgi

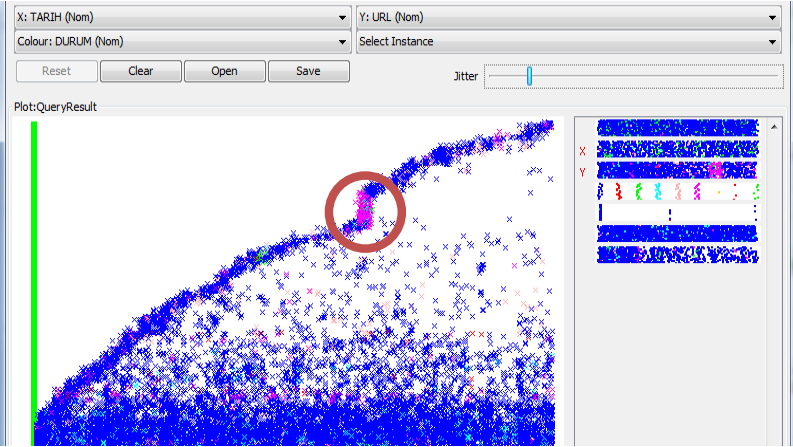
Sadece kırmızı renkle işaretlenen ve kare ile gösterilen küme detaylı incelendiğinde robot olmadığı anlaşılan bir IP numarasından, çok kısa süre içerisinde fazla sayıda istekte bulunulduğu ve istek yapılan URL alanı incelendiğinde saldırı denemelerinde bulunulduğu tespit edilmiştir. Şekil 5’de görüldüğü gibi istek herhangi bir web robotundan gelmemektedir. Arka arkaya aynı IP ‘den gelmektedir. İstek yapılan URL incelendiğinde saldırı girişiminde bulunulduğu görülmektedir.



Şekil 4.5. Saldırı girişimleri içeren istekler

**4.3.2 Önişleme-2 verisine göre analiz**

Önişleme-1 verisine kümeleme ve birliktelik algoritmaları uygulandığında web robotlarının kayıtlarının çok fazla olmasından dolayı; çıkan sonuçlar kullanıcı, saldırı vb. analizlerinde yanılmalara sebep olabilmektedir. Web robotlarının yer aldığı log satırları temizlenerek elde edilen Önişleme-2 verisine kümeleme ve birliktelik algoritmaları uygulanarak anlamlı sonuçlar elde edilmiştir. WEKA’nın Vizualize sekmesinden tarih, URL ve durum alanları için yayılım dağılımına bakıldığında Şekil 4.6’da görüldüğü gibi sadece 1 küme oluştuğu ve bu kümenin de daha önce önişleme-1 verisinde saldırı girişiminde bulunulduğu tespit edilen kümeyle aynı küme olduğu görülmektedir



Şekil 4.6. Önişleme-2 verisi yayılım dağılımı

Böylece Önişleme-2 verisinde web robotlarının temizlenmesiyle daha etkin sonuç elde edilmiştir. Yine aynı veriye kümeleme algoritmalarından Kmeans; URL, tarih ve durum alanları için uygulanmıştır. Şekil 4. 7’de elde edilen sonuç sunulmuştur.

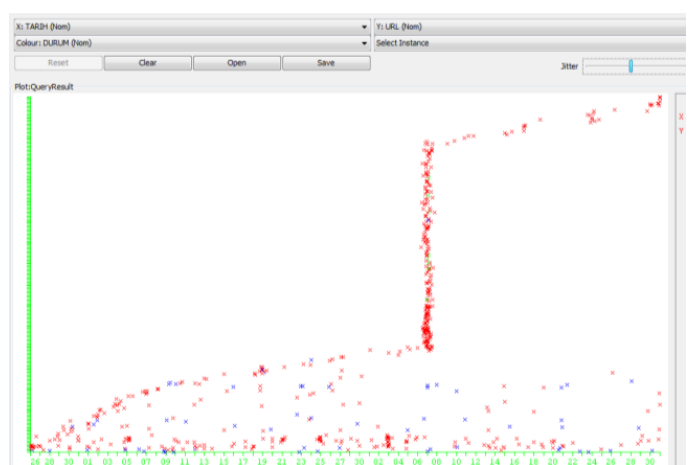
 Şekil 4.7. K-means kümeleme algoritması ile yapılan analiz

Burada başarılı (200 durum kodu) 2 istek için küme oluştuğu görülmektedir ve istek yapılan sayfaların da en çok ziyaret edilen sayfalardan oldukları bilinmektedir. 404 hata durum koduyla ilgili Visualize sekmesinde yayılım dağılımında küme oluştuğu halde burada Kmeans kümeleme algoritmasıyla oluşmamıştır. Çünkü web loglarında genel olarak, özellikle geniş tarih aralığında çalışma yapıldığı durumlarda, 200 serili durum kodlarının yer aldığı istek sayısına göre 400 serili hata durum kodlarının yer aldığı istek sayısı oldukça azdır. Dolayısıyla özellikle saldırı tespitine yönelik analiz yapılacağında 400 serili hata durum kodları haricindeki kayıtların önişleme aşamasında silinmesi gerekmektedir.

Birliktelik kuralları konusundaki en meşhur algoritmalardan biri Apriori algoritmasıdır [44]. Önişleme-2 verisinde tarih, URL ve durum alanları için en az destek değeri 0.1 ve en az güven değeri 0.9 için Apriori algoritmasıyla başarılı (200 durum kodu) ve en sık ziyaret edilen sayfalar arasında birliktelik kuralları oluştuğu görülmüştür.

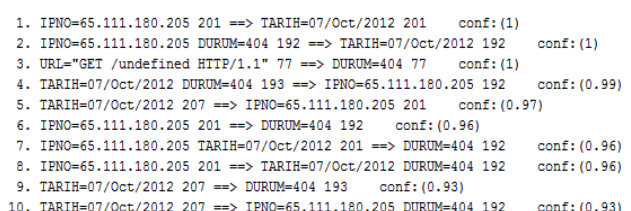
**4.3.3 WEKA ile saldırı Tespiti**

Bir önceki bölümde WEKA ile istatistiksel analiz yapılırken kümeleme ve birliktelik kurallarına göre herhangi bir saldırı girişimi tespit edilememiştir. Sadece WEKA’nın Visualize sekmesinden saldırı girişiminde bulunulduğu tahmin edilmiştir. Bu bölümde doğrudan saldırı tespitine yönelik sadece 400 serili durum kodlarının bulunduğu log satırlarını içeren önişleme-3 verisi kullanılarak analiz yapılmıştır. Veri WEKA’ya yüklendikten sonra önişleme sekmesinde, aynı IP’den çok fazla sayıda hatalı istek alındığı görülmüştür. Şekil 4.8. ’de visualize sekmesinde ise tarih, URL ve durum alanları için yayılım dağılımı incelendiğinde aynı tarihte çok farklı URL’lere istek yapıldığı ve çok sayıda hata durum kodu alındığı görülmektedir. Bu alandaki verilerin detaylı bilgileri incelendiğinde saldırı girişiminde bulunulduğu açıkça görülmektedir.



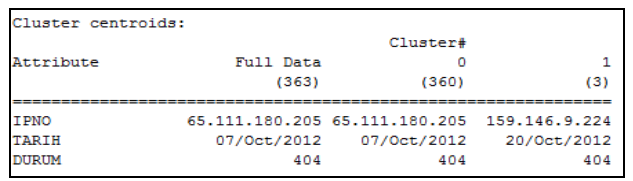
Şekil 4.8. Önişleme-3 verisinin visualize ekranından analizi

Önişleme-3 verisine Apriori algoritmasını çalıştırarak en az destek değeri 0.1 ve en az güven değeri 0.9 için birliktelik kurallarını uyguladığımızda Şekil 9’daki sonuç elde edilmiştir. Buradaki sonuçlara göre; 65.111.180.205 201 numaralı IP numarasının 07 Ekim 2012 tarihinde sayfayı ziyaret etmiş olduğu ve ilk satırda 201 kural oluşurken ikinci satırda durum kodu 404 olan istekler için 192 satır oluştuğu görülmektedir. Yani aynı IP’den gönderilen isteklerin hepsi ilk kez 07 Ekim 2012 tarihinde gönderilmiş ve neredeyse isteklerin tamamında 404 ile sayfa görüntülenememiştir.



Şekil 4.9. Apriori algoritması sonucunda oluşturulan

Tarih, IP numarası ve durum alanları için Kmeans kümeleme algoritması çalıştırılmış ve iki küme oluştuğu görülmüştür. Şekil 10’da elde edilen sonuç sunulmuştur. Kümelerden biri daha önce saldırı girişiminde bulunduğu tespit edilen IP numarasını içermektedir. Diğer küme ise çok az sayıda veri içermektedir.



Şekil 10. K-means ile elde edilen kümeler

Bu bölümde WEKA ile saldırı girişiminde bulunduğu kesin olarak tespit edilmiştir. Bu analizlerden elde edilen saldırı girişiminin tespit edildiği tarih, durum kodu, URL gibi bilgiler web logları saldırı analiz aracı olan Apache Scalp ile saldırı türünü ve saldırının başarılı olup olmadığını tespit etmek üzere kullanılmıştır.

# SONUÇLAR VE ÖNERİLER

Gerçekleştirilen istatistiksel analizler neticesinde kullanıcı eğilimleri ve genel site kullanımı hakkında elde edilen bulgularla site geliştiricilerine web sitesinin iyileştirilmesine ve geliştirilmesine yönelik katkıda bulunulacağı düşünülmektedir. WEKA’da gerçekleştirilen analizlerde hem web robotlarıyla hem de web robotları olmadan analiz gerçekleştirilerek ziyaretçi sayısı, başarılı istek sayısı, hatalı istek sayısı gibi sonuçlar elde edilmiştir. Kümeleme ve birliktelik kuralları ile belirli tarihlerde durum kodlarına göre en çok ziyaret edilen sayfalar belirlenmiştir.

Elde edilen sonuçlar ve öneriler aşağıda sunulmuştur:

Genel olarak sayfalara yapılan ziyaretler, sayfa görüntüleme ve bant genişliği hakkında istatistiksel bilgiler verilmiştir. Sayfaya yapılan ziyaretlerde web robotlarının gönderdiği isteklerle kullanıcı isteklerinin birbirine çok yakın sayıda olduğu görülmüştür. Özellikle kullanıcılarla ilgili yapılacak analizlerde yanıltıcı olmaması için web robotlarının kayıtlarının temizlenmesi gerektiği sonucuna varılmıştır.

Web robotlarının temizlenmiş olduğu önişleme-2 verisi web madenciliği ile analiz edildiğinde 200 durum koduyla en çok ziyaret edilen sayfalar tespit edilmiştir. Web loglarında geniş tarih aralığında işlem yapıldığında çoğu zaman işlemin başarılı olduğuna dair durum kodu alınacağından visualize ekranında oluşan saldırı örüntüleri kümeleme veya birliktelik kuralları ile görülememiştir. Bu nedenle web madencilliğinde özellikle saldırı tespitine yönelik analiz gerçekleştirileceğinde 400 serili hata durum kodlarının haricindeki log satırlarının temizlenmesi gerekmektedir. Bu şekilde elde edilen önişlem-3 verisine web madenciliği algoritmaları uygulandığında doğrudan saldırı tespiti yapılmıştır.

Bir sonraki çalışamada, web loglarından saldırı tespit işleminin tam otomatize gerçekleştirilmesi düşünülmektedir.

**KAYNAKLAR**

[1] A. Adamov, “Data mining and analysis in depth. case study of

Qafqaz University HTTP server log analysis”, In Application of

Information and Communication Technologies (AICT) IEEE

8th International Conference, 1-4, 2014.

[2] A. S. Yumnam, Y. Chaitanya Sreeram & S. A. Naeem,

“Overview: Weblog mining, privacy issues and application of Web

Log mining”, In Computing for Sustainable Global Development

(INDIACom) International Conference, 638-641, 2014.

[3] A. D. Khairkar, D. D. Kshirsagar & S. Kumar, “Ontology for

Detection of Web Attacks”, In Communication Systems and

Network Technologies (CSNT) International Conference, 612-

615, 2013.

[4] B. Mobasher, R. Cooley, J. Srivastava, “Automatic Personalization

based on Web Usage Mining”, Communications of the ACM,

43(8),142-151, 2000.

[5] A. Vahaplar, M. M. İnceoğlu, “Veri Madenciliği ve Elektronik

Ticaret”, Türkiye’de Internet Konferansları VII, 2001.

[6] R. Daş, Web Kullanıcı Erişim Kütüklerinden Bilgi Çıkarımı,

Doktora Tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Elazığ,

2008.

[7] J. P. Leite, “Analysis of log files as a security aid”, In

Information Systems and Technologies (CISTI) IEEE 6th

Iberian Conference, 1-6, 2011.

[8] M. Nagappan & M. A. Vouk, “Abstracting log lines to log event

types for mining software system logs”, In Mining Software

Repositories 7th IEEE Working Conference, 114-117, 2010.

[9] K.C. Burçak, Kırıkkale Üniversitesi Web Sitesinin Kullanıcı

Örüntülerinin Web Madenciliği ile Analizi, Yüksek Lisans Tezi,

Kırıkkale Üniversitesi Fen Bilimleri Enstitüsü, Kırıkkale, 2012.

[10] S. S. Vernekar & A. Buchade, “MapReduce based log file analysis

for system threats and problem identification”, In Advance

Computing Conference (IACC), 831-835, 2013.

[11] O. Ozulku, N. F. Fadhel, D. Argles & G. B. Wills, “Anomaly

detection system: Towards a framework for enterprise log

management of security services”, In Internet Security

(WorldCIS) 2014 World Congress, 97-102, 2014.

[12] İ. Haberal, Veri Madenciliği Algoritmaları Kullanılarak Web

Günlük Erişimlerinin Analizi, Yüksek Lisans Tezi, Başkent

Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2007.

[13] T. Hussain, S. Asghar & N. Masood, “Web usage mining: A

survey on preprocessing of web log file”, In International

Conference on Information and Emerging Technologies

(ICIET), 1-6, 2010.

[14] B. Özakar & H. Püskülcü, “Web içerik ve web kullanım

madenciliği tekniklerinin entegrasyonu ile oluşmuş bir veri

tabanından nasıl yararlanılabilir?”, 2002.

[15] D. S. Sisodia & S. Verma, “Web usage pattern analysis through

web logs: A review”, In Computer Science and Software

Engineering (JCSSE) International Joint Conference, 49-53,

2012.

[16] R. Yevale, M. Dhage, T. Nalawade & T. Kaule, “Unauthorized

Terror Attack Tracking Using Web Usage Mining”, (IJCSIT)

International Journal of Computer Science and Information

Technologies, 5 (2), 1210-1212, 2014.

[17] T. Gržinic, , T. Kišasondi, J. Šaban, “Detecting anomalous Web

server usage through mining access logs”, Central European

Conference on Information and Intelligent Systems, 228-296,

2013.

[18] M. Turan, Web Mining: Pattern Discovery On The World Wide

Web, Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi Fen Bilimleri

Enstitüsü, İzmir, 2011

[19] A. Ashoor, S. Gore , “Intrusion Detection System (IDS):

Case Study”, International Conference on Advanced

Materials Engineering,2011.

[20] J. Stenico, L. Ling, “Network Traffic Monitoring and

Analysis”, The State of the Art in Intrusion Prevention and

Detection, 23-46, 2014.

[21] The KDD CUP 1999 Data. 1999,

http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

(Erişim zamanı; Nisan, 2, 2018)

[22] NSL-KDD, URL: http:// unb.ca/cic/datasets/nsl.html

(Erişim zamanı; Nisan, 3, 2018)

[23] M. Hall, E. Frank, J. Holmes, B. Pfahringer, P.

Reutemann, and I. Witten, “The WEKA data mining

software: An update,” ACM SIGKDD Explor. Newslett.,

vol. 11, no. 1, pp. 10–18, 2009.

[24] R Language Definition, R Core Team, URL:

ftp://155.232.191.133/cran/doc/manuals/r-devel/R-lang.pdf

(Erişim zamanı; Nisan, 3, 2018)

**EKLER**

**EK A:** Karar kuralları

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| rule 1. | A1 = 2 | A2 = 3 | A3 = 3 | A7 = 2 | A9 = 1 | Dec = 1 |  |  |  |
| rule 2. | A6 = 3 | A7 = 2 | Dec = 1 |  |  |  |  |  |  |
| rule 3. | A2 = 3 | A3 = 2 | A5 = 2 | A6 = 1 | A7 = 2 | A9 = 1 | Dec = 1 |  |  |
| rule 4. | A2 = 2 | A3 = 3 | A4 = 2 | A5 = 2 | A7 = 2 | Dec = 1 |  |  |  |
| rule 5. | A2 = 3 | A3 = 3 | A6 = 2 | A7 = 2 | Dec = 1 |  |  |  |  |
| rule 6. | A3 = 3 | A4 = 3 | A6 = 1 | A7 = 2 | Dec = 1 |  |  |  |  |
| rule 7. | A1 = 1 | A2 = 3 | A4 = 2 | A7 = 2 | A9 = 1 | Dec = 1 |  |  |  |
| rule 8. | A2 = 3 | A6 = 3 | A7 = 1 | A8 = 2 | A9 = 1 | Dec = 1 |  |  |  |
| rule 9. | A1 = 1 | A2 = 2 | A6 = 2 | A7 = 2 | A8 = 1 | Dec = 1 |  |  |  |
| rule 10. | A3 = 1 | A4 = 1 | A5 = 2 | A7 = 2 | A8 = 3 | A9 = 1 | Dec = 1 |  |  |
| rule 11. | A2 = 2 | A5 = 2 | A6 = 2 | A7 = 3 | A8 = 2 | A9 = 2 | Dec = 1 |  |  |
| rule 12. | A1 = 1 | A2 = 1 | A4 = 1 | A6 = 3 | A8 = 2 | A9 = 2 | Dec = 1 |  |  |
| rule 13. | A1 = 1 | A4 = 3 | A6 = 2 | A7 = 2 | A8 = 1 | Dec = 1 |  |  |  |
| rule 14. | A1 = 2 | A2 = 3 | A4 = 1 | A6 = 3 | A8 = 3 | A9 = 1 | Dec = 1 |  |  |
| rule 15. | A1 = 1 | A3 = 2 | A4 = 3 | A5 = 2 | A7 = 2 | A8 = 1 | A9 = 1 | Dec = 1 |  |
| rule 16. | A3 = 3 | A6 = 2 | A7 = 2 | A9 = 2 | Dec = 1 |  |  |  |  |
| rule 17. | A2 = 1 | A4 = 2 | A7 = 3 | A8 = 2 | A9 = 1 | Dec = 1 |  |  |  |
| rule 18. | A1 = 1 | A2 = 3 | A6 = 2 | A7 = 3 | A8 = 2 | A9 = 2 | Dec = 1 |  |  |
| rule 19. | A3 = 3 | A6 = 5 | A7 = 1 | A9 = 1 | Dec = 1 |  |  |  |  |
| rule 20. | A2 = 3 | A4 = 3 | A7 = 2 | A9 = 2 | Dec = 1 |  |  |  |  |
| rule 21. | A1 = 2 | A3 = 3 | A6 = 2 | A7 = 2 | Dec = 1 |  |  |  |  |
| rule 22. | A1 = 1 | A2 = 1 | A3 = 1 | A4 = 2 | A6 = 3 | A7 = 3 | A8 = 3 | Dec = 1 |  |
| rule 23. | A1 = 1 | A2 = 3 | A3 = 1 | A6 = 4 | A8 = 1 | A9 = 1 | Dec = 1 |  |  |
| rule 24. | A2 = 2 | A3 = 2 | A6 = 4 | A7 = 1 | A9 = 2 | Dec = 1 |  |  |  |
| rule 25. | A2 = 3 | A3 = 2 | A6 = 2 | A7 = 3 | A8 = 3 | A9 = 2 | Dec = 1 |  |  |
| rule 26. | A1 = 1 | A2 = 2 | A3 = 2 | A4 = 2 | A5 = 2 | A6 = 1 | A7 = 1 | A8 = 1 | Dec = 1 |
| rule 27. | A2 = 3 | A6 = 3 | A7 = 3 | A8 = 2 | A9 = 2 | Dec = 1 |  |  |  |
| rule 28. | A1 = 1 | A3 = 2 | A4 = 3 | A5 = 2 | A6 = 2 | A8 = 1 | A9 = 1 | Dec = 1 |  |
| rule 29. | A2 = 1 | A4 = 2 | A5 = 2 | A6 = 1 | A7 = 2 | A9 = 2 | Dec = 1 |  |  |
| rule 30. | A2 = 2 | A3 = 2 | A4 = 2 | A7 = 2 | Dec = 1 |  |  |  |  |
| rule 31. | A1 = 1 | A2 = 1 | A5 = 2 | A6 = 3 | A7 = 3 | A8 = 2 | Dec = 1 |  |  |
| rule 32. | A1 = 2 | A2 = 1 | A7 = 2 | A8 = 3 | Dec = 1 |  |  |  |  |
| rule 33. | A2 = 2 | A4 = 3 | A5 = 2 | A7 = 3 | A8 = 2 | A9 = 2 | Dec = 1 |  |  |
| rule 34. | A1 = 1 | A2 = 3 | A4 = 3 | A6 = 3 | A9 = 1 | Dec = 1 |  |  |  |
| rule 35. | A1 = 1 | A2 = 2 | A3 = 1 | A4 = 2 | A6 = 2 | A7 = 1 | A8 = 3 | A9 = 2 | Dec = 1 |
| rule 36. | A6 = 5 | A7 = 2 | Dec = 1 |  |  |  |  |  |  |
| rule 37. | A1 = 2 | A2 = 2 | A3 = 3 | A6 = 3 | A9 = 1 | Dec = 1 |  |  |  |
| rule 38. | A1 = 2 | A2 = 1 | A6 = 2 | A7 = 3 | A8 = 2 | A9 = 2 | Dec = 1 |  |  |
| rule 39. | A3 = 3 | A4 = 2 | A6 = 2 | A7 = 3 | A8 = 1 | A9 = 2 | Dec = 1 |  |  |
| rule 40. | A1 = 1 | A2 = 3 | A3 = 2 | A4 = 1 | A5 = 2 | A6 = 2 | A7 = 3 | A9 = 2 | Dec = 1 |
| rule 41. | A2 = 3 | A4 = 2 | A7 = 3 | A8 = 2 | A9 = 1 | Dec = 1 |  |  |  |
| rule 42. | A1 = 2 | A2 = 1 | A3 = 1 | A4 = 1 | A6 = 3 | A7 = 3 | A8 = 3 | A9 = 2 | Dec = 1 |
| rule 43. | A1 = 2 | A2 = 2 | A3 = 1 | A4 = 3 | A5 = 2 | A6 = 1 | A8 = 3 | A9 = 1 | Dec = 1 |
| rule 44. | A1 = 1 | A2 = 3 | A3 = 1 | A4 = 3 | A6 = 2 | A8 = 3 | Dec = 1 |  |  |
| rule 45. | A2 = 2 | A3 = 3 | A4 = 2 | A5 = 2 | A7 = 3 | A8 = 1 | A9 = 1 | Dec = 1 |  |
| rule 46. | A2 = 2 | A3 = 1 | A4 = 1 | A7 = 2 | A9 = 2 | Dec = 1 |  |  |  |
| rule 47. | A1 = 1 | A2 = 2 | A5 = 1 | A8 = 2 | A9 = 1 | Dec = 1 |  |  |  |
| rule 48. | A2 = 3 | A6 = 5 | A7 = 3 | A8 = 1 | A9 = 1 | Dec = 1 |  |  |  |
| rule 49. | A2 = 3 | A4 = 1 | A6 = 3 | A7 = 3 | A8 = 3 | A9 = 2 | Dec = 1 |  |  |
| rule 50. | A1 = 2 | A2 = 2 | A3 = 1 | A6 = 4 | A9 = 1 | Dec = 1 |  |  |  |
| rule 51. | A2 = 2 | A3 = 2 | A4 = 2 | A6 = 2 | A7 = 3 | A8 = 3 | Dec = 1 |  |  |
| rule 52. | A2 = 2 | A6 = 4 | A7 = 3 | A9 = 1 | Dec = 1 |  |  |  |  |
| rule 53. | A1 = 1 | A4 = 2 | A6 = 2 | A7 = 1 | A8 = 2 | Dec = 1 |  |  |  |
| rule 54. | A1 = 2 | A3 = 2 | A4 = 1 | A6 = 2 | A7 = 2 | Dec = 1 |  |  |  |
| rule 55. | A1 = 2 | A3 = 1 | A4 = 2 | A6 = 2 | A7 = 3 | A8 = 3 | A9 = 2 | Dec = 1 |  |
| rule 56. | A1 = 1 | A2 = 3 | A4 = 3 | A6 = 3 | A7 = 1 | Dec = 1 |  |  |  |
| rule 57. | A1 = 2 | A2 = 2 | A3 = 1 | A4 = 3 | A5 = 2 | A6 = 2 | A8 = 1 | A9 = 1 | Dec = 1 |
| rule 58. | A1 = 1 | A2 = 2 | A3 = 1 | A4 = 1 | A5 = 1 | A6 = 2 | A7 = 3 | A9 = 1 | Dec = 1 |
| rule 59. | A1 = 1 | A2 = 2 | A3 = 2 | A7 = 2 | A9 = 2 | Dec = 1 |  |  |  |
| rule 60. | A4 = 2 | A6 = 3 | A7 = 3 | A8 = 3 | A9 = 1 | Dec = 1 |  |  |  |
| rule 61. | A1 = 2 | A3 = 1 | A5 = 2 | A6 = 5 | A7 = 3 | A8 = 3 | Dec = 1 |  |  |
| rule 62. | A2 = 3 | A3 = 1 | A4 = 3 | A5 = 1 | A7 = 2 | Dec = 1 |  |  |  |
| rule 63. | A1 = 1 | A2 = 2 | A6 = 4 | A8 = 3 | A9 = 1 | Dec = 1 |  |  |  |
| rule 64. | A2 = 3 | A4 = 3 | A6 = 4 | Dec = 1 |  |  |  |  |  |

**ÖZGEÇMİŞ**

2000 yılı Mayıs ayının 3’ünde Isparta'da doğdu. 2019 yılında Sakarya Üniversitesine yatay geçiş ile geldi. Üniversite yıllarında Siber Güvenlik alanında yoğunlaşmış olup bu alanda 2 staj tamamladı. Çalışmalarına özel sektörde devam etmektedir.

**BSM 401 BİLGİSAYAR MÜHENDİSLİĞİ TASARIMI**

**Değerlendİrme ve Sözlü Sınav Tutanağı**

KONU :

**Değerlendirme ve Sözlü Sınav Tutanağı:**Bilgisayar Mühendisliği Tasarımı konu başlığı ve ilgili öğrencilerin bilgileri girildikten sonra alt kısma da jüri adına danışman hocanın ismi yazılmalıdır. Bu form BSM401 Bilgisayar Mühendisliği Tasarımı Kitapçığınızın son sayfasıdır. Danışman hocanız bu formu doldurup imzalayacaktır.

**Burayı çıktı almadan önce siliniz.**

ÖĞRENCİLER (Öğrenci No/AD/SOYAD):

|  |  |  |  |
| --- | --- | --- | --- |
| Değerlendirme Konusu | İstenenler | Not Aralığı | Not |
| **Yazılı Çalışma** |  |  |  |
| **Çalışma klavuza uygun olarak hazırlanmış mı?** | x | 0-5 |  |
| **Teknik Yönden** |  |  |  |
| **Problemin tanımı yapılmış mı?** | x | 0-5 |  |
| Geliştirilecek yazılımın/donanımın mimarisini içeren blok şeması (yazılımlar için veri akış şeması (dfd) da olabilir) çizilerek açıklanmış mı? |  |  |  |
| Blok şemadaki birimler arasındaki bilgi akışına ait model/gösterim var mı? |  |  |  |
| Yazılımın gereksinim listesi oluşturulmuş mu? |  |  |  |
| Kullanılan/kullanılması düşünülen araçlar/teknolojiler anlatılmış mı? |  |  |  |
| Donanımların programlanması/konfigürasyonu için yazılım gereksinimleri belirtilmiş mi? |  |  |  |
| UML ile modelleme yapılmış mı? |  |  |  |
| Veritabanları kullanılmış ise kavramsal model çıkarılmış mı? (Varlık ilişki modeli, noSQL kavramsal modelleri v.b.) |  |  |  |
| Projeye yönelik iş-zaman çizelgesi çıkarılarak maliyet analizi yapılmış mı? |  |  |  |
| Donanım bileşenlerinin maliyet analizi (prototip-adetli seri üretim vb.) çıkarılmış mı? |  |  |  |
| Donanım için gerekli enerji analizi (minimum-uyku-aktif-maksimum) yapılmış mı? |  |  |  |
| Grup çalışmalarında grup üyelerinin görev tanımları verilmiş mi (iş-zaman çizelgesinde belirtilebilir)? |  |  |  |
| Sürüm denetim sistemi (Version Control System; Git, Subversion v.s.) kullanılmış mı? |  |  |  |
| Sistemin genel testi için uygulanan metotlar ve iyileştirme süreçlerinin dökümü verilmiş mi? |  |  |  |
| Yazılımın sızma testi yapılmış mı? |  |  |  |
| Performans testi yapılmış mı? |  |  |  |
| Tasarımın uygulamasında ortaya çıkan uyumsuzluklar ve aksaklıklar belirtilerek çözüm yöntemleri tartışılmış mı? |  |  |  |
| **Yapılan işlerin zorluk derecesi?** | x | 0-25 |  |
| **Sözlü Sınav** |  |  |  |
| **Yapılan sunum başarılı mı?** | x | 0-5 |  |
| **Soruları yanıtlama yetkinliği?** | x | 0-20 |  |
| **Devam Durumu** |  |  |  |
| **Öğrenci dönem içerisindeki raporlarını düzenli olarak hazırladı mı?** | x | 0-5 |  |
| **Diğer Maddeler** |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **Toplam** |  |  |  |

Danışman : Doç. Dr. Nilüfer Yurtay

danışman imzası: