Avrupa Bilim ve Teknoloji Dergisi
Sayı 20, S. 498-505, Aralık 2020

Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. 20, pp. 498-505, December 2020

Research Article

Birliktelik Kural Analizi Tabanlı İzleme ve Bayes Ağları ile Operasyonel Teknoloji Sistemlerinde Siber Güvenlik Analizi

Volkan Altuntaş[1*]

[1*] Bursa Teknik Üniversitesi, Rektörlük, Bursa, Türkiye, (ORCID: 0000-0003-3144-8724), volkan.altuntas@btu.edu.tr

(İlk Geliş Tarihi 15 Eylül 2020 ve Kabul Tarihi 27 Ekim 2020)

(**DOI:** 10.31590/ejosat.800954)

ATIF/REFERENCE: Altuntas, V. (2020). Birliktelik Kural Analizi Tabanlı İzleme ve Bayes Ağları ile Operasyonel Teknoloji Sistemlerinde Siber Güvenlik Analizi. *Avrupa Bilim ve Teknoloji Dergisi*, (20), 498-505.

Öz

Artan teknoloji kullanımı ve sistemler arası entegrasyonun getirdiği kullanım kolaylığına ek olarak oluşan güvenlik açıkları günümüzde siber güvenliğin önemini artırmaktadır. Hükümetler, askeri, kurumsal, finans veya tıbbi kuruluşlar güvenlik tehdidi altında bulunmaktadır. Yaygınlaşan teknolojiyle beraber güvenli olmayan bilgisayar sistemleri, yıkıcı kesintilere, hassas bilgilerin ifşasına ve sahtekârlıklara yol açabilir. Endüstriyel varlıkların izlendiği ve kontrol edildiği sistemler Operasyonel Teknoloji (OT) olarak adlandırılır. Denetleyici Kontrol (PLC) ve Veri Toplama (SCADA) sistemleri OT sistem örneklerindendir. Bu sistemlere yapılan saldırılar, hayati önem taşıyan temel hizmetlerin sunulmasını engelleyebilir, ciddi ekonomik veya sosyal sonuçlara yol açabilir. Yapılarına güvenlik mekanizmalarının uygulanamayışı veya sahip oldukları eski bilişim teknolojilerinin yenilememesi sebepleri ile OT sistemler siber saldırılara karşı en savunmasız sistemlerdir. Bu çalışma kapsamında OT sistemleri için saldırı tespit ve uyarı sistemi geliştirilmiştir. OT sistemlerin mevcut durumu göz önüne alınarak tasarlanan sistem, veri erişimini OPC sunucu üzerinden yapmakta ve bu sayede tüm OT sistemlerine minimum değişiklik ile bağlantı sağlayabilecek yapıdadır. OT sistemde meydana gelen işlemlerin algılanması ve analiz edilebilmesi için birliktelik analizi tabanlı aktivite kayıt oluşturma algoritması geliştirilmiş. Geliştirilen bu algoritma ile OT süreç bilgisi olmadan tüm aktivitelerin yorumlanabilmesine olanak sağlanmıştır. Oluşan aktivite bilgilerinin analizi için Bayes ağ tabanlı bir öğrenme sistemi tasarlanmıştır. Bu sistem sayesinde elde edilen kayıtlardan Bayes ağları oluşturulmakta, sistem çalışma anında oluşan tüm aktiviteler değerlendirilerek olasılıklarına göre "Güvenli", "Riskli" ve "Siber Saldırı" olarak gruplandırılmakta ve OT yetkililerine sunulmaktadır. Önerilen sistem her tür OT aktivitesine adapte olabilecek yapıdadır. Sistem mimari yapısı gereği sadece siber saldırı kategorisindeki işlemleri tespit etmekte kalmayıp sistemde yetkili kişilerin hatalı yâda kasıtlı müdahale veya program değişimlerini de algılayabilmektedir. Deneysel çalışmalarımız OT sisteme yapılan her türlü saldırının tespit edilebildiğini göstermektedir. Bu çalışma OT davranışlarının modellenerek öğrenildiği ve anormal davranışların tespit edilerek siber saldırıların tespit edildiği ilk çalışmadır.

Anahtar Kelimeler: Siber Güvenlik, Operasyonel Teknolojiler, Akıllı Sistemler, Makine Öğrenmesi.

Cyber Security Analysis at Operational Technology Systems with Association Rule-Based Monitoring and Bayesian Networks

Abstract

In addition to the ease of use brought by the increasing use of technology and integration between systems, the resulting security vulnerabilities increase the importance of cybersecurity today. Governments, military, corporate, financial, or medical organizations are under security threats. With the spread of technology, unsafe computing systems can lead to devastating interruptions, sensitive information disclosure, and fraud. Systems in which industrial assets are monitored and controlled are called Operational Technology (OT). Supervisory Control (PLC) and Data Acquisition (SCADA) systems are examples of OT systems. Attacks on these systems can hinder the delivery of vital essential services and have serious economic or social consequences. OT systems are the most vulnerable systems against cyber-attacks since security mechanisms cannot be applied to their structures or the old information technologies they have are not renewed. Within the scope of this study, an intrusion detection and warning system has been developed for OT systems. The system, which is designed considering the current situation of OT systems, makes data access through the OPC server and thus, it

* Sorumlu Yazar: volkan.altuntas@btu.edu.tr

is in a structure that can provide a connection to all OT systems with minimum changes. To detect and analyze the transactions occurring in the OT system, an association analysis based activity record creation algorithm has been developed. With this developed algorithm, it is possible to interpret all activities without OT process knowledge. A Bayes network-based learning system was designed to analyze the activity information. Bayes networks are created from the records obtained through this system and all activities that occur during the system operation are evaluated, grouped as "Safe", "Risky" and "Cyber Attack" according to their probabilities and presented to OT officials. The proposed system is capable of adapting to all types of OT activities. Due to the architecture of the system, it not only detects transactions in the cyberattack category but also detects faulty or deliberate intervention or program changes by authorized persons in the system. Our experimental studies show that any attack on the OT system can be detected. This study is the first study in which OT behaviors are modeled and learned, and cyberattacks are detected by detecting abnormal behaviors.

Keywords: Cyber Security, Operational Technologies, Intelligent Systems, Machine Learning.

1. Giriş

Güvenli olmayan bilgisayar sistemleri, yıkıcı kesintilere, hassas bilgilerin ifşasına ve sahtekârlıklara yol açabilir. Bilgisayarları, ağları, programları ve verileri saldırılardan, yetkisiz erişim veya değişimlerden veya kullanım dışı kalmasından korumak için tasarlanmış teknolojiler ve süreçler kümesi Siber güvenlik olarak tanımlanır. Siber güvenlik sistemleri, ağ güvenlik sistemleri ve bilgisayar güvenlik sistemlerinden oluşur. Bu sistemler güvenlik duvarı, anti virüs yazılımı ve saldırı tespit sistemi içerebilir. Saldırı tespit sistemleri, bilgi sistemlerinin yetkisiz kullanımının, çoğaltılmasının, değiştirilmesinin ve imhasının tespit edilmesine, belirlenmesine ve tanımlanmasına yardımcı olur [1]. Güvenlik ihlalleri, kuruluş dışından gelen saldırılar ve kuruluş içinden saldırılar olarak ikiye ayrılabilir. Kötüye kullanım tabanlı, anormallik tabanlı ve karma olmak üzere saldırı tespit sistemlerinin kullandığı üç ana siber analiz türü vardır. Kötüye kullanıma dayalı teknikler, bu saldırıların imzalarını kullanarak bilinen saldırıları tespit etmek için tasarlanmıştır. Çok fazla sayıda yanlış alarm oluşturmadan bilinen saldırı türlerini tespit etmek için etkilidirler. Veri tabanının kurallar ve imzalarla sık sık manuel olarak güncellenmesini gerektirirler. Kötüye kullanıma dayalı teknikler sıfırıncı gün saldırıları tespit edemez. Anormalliğe dayalı teknikler, normal ağ ve sistem davranışını modeller ve anormallikleri normal davranıştan sapmalar olarak tanımlar. Sıfırıncı gün saldırılarını tespit etme kabiliyetleri nedeniyle önemlidir. Diğer bir avantaj, normal etkinlik profillerinin her sistem, uygulama veya ağ için özelleştirilmesidir. Bu nedenle saldırganların tespit edilmeden hangi etkinlikleri gerçekleştirebileceklerini bilmelerini zorlaştırır. Anormallik temelli tekniklerin oluşturduğu uyarılar yeni saldırı imzalarının oluşturulması için kullanılabilir. Anormalliğe dayalı tekniklerin temel dezavantajı, yüksek yanlış alarm oranlarıdır. Önceden görülmemiş tüm sistem davranışları anormallik olarak kategorize edilebilir. Hibrit teknikler, kötüye kullanım ve anormallik tespitini birleştirir. Bilinen izinsiz girişlerin tespit oranlarını artırmak ve bilinmeyen saldırılarında tespit edilmesine olanak tanımak için kullanılırlar [2].

Operasyonel Teknoloji (OT), endüstriyel proses varlıklarını ve endüstriyel ekipmanları izler ve yönetir. OT sistemleri bina, ulaşım, endüstri gibi sektörlerde elektrikle çalışan ekipmanlar kullanılmaya başlandığından buya yana var olduğu için Bilgi Teknolojisinden (BT) eskiye dayanmaktadır. Denetleyici Kontrol (PLC) ve Veri Toplama (SCADA) sistemleri OT sistem örneklerindendir. OT sistemler endüstriyel sektörlerdeki süreçleri kontrol eder [3]. Bu sektörler her ülke için kritik önemde olup, bu sistemlere yapılacak olası olası saldırılar, temel hizmetlerin aksamasına, ciddi ekonomik veya sosyal sonuçlara yol açacak altyapı varlık kayıtlarına ya da can kayıplarına neden olur [4]. Günümüzde PLC ve SCADA sistemlerinin kullanılmadığı sektör yok denecek kadar az sayıdadır [5]. OT sistemlerin sorunsuz ve güvenilir çalışması, hem veri toplama hem de kontrolün kritik öneme sahip olduğu sektörler için hayati öneme sahiptir. OT sistemlerde yaşanan yaygın veya uzun süreli kesintiler sonuç olarak devlet ve toplumda ciddi rahatsızlıklara neden olabilir [6]. Bir SCADA sisteminin arızalanmasının sonuçları zararlı olabilir ve bir ekipmandan kaynaklanan mali kayıptan insan hayatının kaybına kadar çevresel hasara kadar değişebilir. Genel olarak güvenlik ve özellikle siber güvenlik OT sistemlerinin temel hedeflerinden değildi [7]. Güvenlik fiziksel yalıtım ve ürüne özel iletişim protokolleri ile sağlanmaktaydı. Yıllar boyunca OT sistem güvenliği kullanım şartları kaynaklı erişim kısıtları sebebi ile doğal olarak bulunmaktaydı. Son on yılda sanayinin dijital dönüşümü ile gerçekleşen yenilikler ile OT sistemler izole kullanım ortamlarını kaybetti ve siber güvenliğiyle ilgili bir dizi standart ve direktif ortaya çıktı. Günümüzde OT sistemler ileri teknoloji sistemlerine sahip olmasına karşın artan karmaşıklık, modernizasyon, gerçek zamanlı sürekli çalışma gereksinimi, dağıtılmış ve çok bileşenli mimari gibi sebepler ile OT sistemlere yönelik siber tehditler artmaya devam etmektedir.

İnternetin sürekli artan gücü, birden çok yerden eşzamanlı saldırıları kolaylaştırır. Bir saldırının en yüksek etkisi, bir saldırganın bir OT sisteminin denetleyici kontrol erişimine erişim kazanması ve felaketle sonuçlanabilecek hasarlara neden olabilecek kontrol eylemlerini başlatmasıdır [8]. Son zamanlarda standartlaştırılmış protokolleri kullanma eğilimi ile daha fazla yardımcı program, geniş alan iletişimi için İnternet protokolü IP tabanlı sisteme doğru ilerliyor. Bunun sonucu oluşan entegrasyon, beraberinde yeni güvenlik açıklarını getirmektedir. OT sistemlerinin internete bağlanmasıyla ilgili güvenlik açığı riskleri bilinmektedir [9]. İnternet üzerinden iletişime artan bağımlılık, sorunun önemine ve büyüklüğüne katkıda bulunmuştur. OT sistemlerine ilişkin güvenlik bilinci ve personel eğitimi çok önemlidir [10]. Son araştırmalar, kasıtlı sabotajıda içeren güvenlikte karşılıklı bağımlılık modellemesini, bilgi mimarisi ve iletişim etkileşimindeki iyileştirme gereksinimlerini ortaya koymaktadır [11].

Yaygın olan kapalı endüstriyel ortamda, teknolojinin gelişmesiyle ilgili önemli bilgiler ve patentler ticari markalar için gizli bilgi kategorisindedir. PLC ve SCADA'ların kullanılmaya başlanmasından bu yaya satıcılar endüstriyel ortamı hedefleyen kendi tescilli donanım ve yazılım çözümlerini geliştirmektedir [12]. Son zamanlarda, iletişim ağlarının ve İnternetin gelişmesiyle bu sistemler çok çeşitli siber saldırılara karşı savunmasız hale gelmiştir. Bununla birlikte, PLC ve

SCADA cihazlarının donanımı ve donanım yazılımı hakkında bilgilere erişim kısıtlıdır. Bu sebepler, PLC ve SCADA'lar için siber güvenlik araştırması yapmayı zorlaştırır ve güvenliği sağlanması için mevcut sistemlere güvenmek zorunda bırakır.

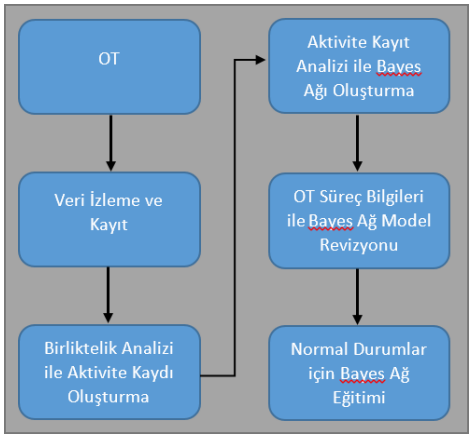
Davis ve arkadaşları [13] yaptıkları çalışma ile iletişim için genel ağların kullanımının güç sistemleri üzerinde oluşturduğu güvelik açıklarını incelemişlerdir. Akıllı şebeke sistemlerinin güç dağıtım sistemleri üzerine olan siber güvenlik unsurları Ericsson ve arkadaşları [14] tarafından araştırılmıştır. Simo ve arakadaşları yapmış oldukları çalışma ile PRIME standardında bulunan siber güvenlik açıklarını ortaya çıkarmışlardır [15]. BCIT Endüstriyel Güvenlik Olay Veri tabanında (ISID) toplanan olay bilgileri Byres ve arkadaşları [16] tarafından özetlenerek operasyonel teknoloji sistemlerini doğrudan etkileyen olaylar açıklanmıştır. Xie ve arkadaşları [17] yaptıkları çalışma ile siber güvenlikteki belirsiz yönlerin modellenmesini ve kurumsal ağlarda güvenlik analizi için kullanılmasını sağlamışlardır.

Mevcut durum göz önüne alındığında, tüm OT sistemlere uygulanabilecek, mevcut çalışma ortamlarını etkilemeyecek, hedef OT sistem hakkında minimum bilgi ve minimum değişiklik ile siber güvenlik tehlikelerine karşı savunma, tespit ve önlem sağlayabilecek tekniğin anormallik tespit yöntemi olduğu görülmektedir. Anormallik tespiti, veri madenciliği çalışmalarının önemli tekniklerinden biridir, dolandırıcılık, siber güvenlik ve arıza tespitinde kullanılmaktadır [18]. Veri analizi çalışmalarına sisteme ait log verileri etkin bir şekilde kullanılabilmektedir [19]. Günlük verileri, kaynaklarının güvenlik ve karar mekanizmalarına ait temel referans bilgileri içerdikleri için erken tespit ve erken uyarı sistemleri için vazgeçilmez kabul edilir [20].

OT sistemlerinin mevcut durumu ve siber güvenlik risklerinin göz önüne alındığı bu çalışmada, mevcut literatürden farklı olarak, OT sistemleri için birliktelik analizi tabanlı aktivite kayıt oluşturma algoritması geliştirilmiş, aktivite log verilerinin sistematik olarak Bayes ağları ile analiz edilerek anormal durumların tespit edilmesi ve OT sistem yöneticilerinin uyarılması için yeni sistem önerilmiştir. Önerilen sistem her tür OT sistem aktivitesine adapte olabilecek yapıdadır. Entegrasyon için OT sistem müdahalesi gerektirmemekte, OT sistem verilerine erişim sağlayacak bir OPC (Ole for Process Control) sunucu yâda PLC veya SCADA sistem verilerine okuma erişimi yeterli olmaktadır.

2. Materyal ve Metot

OT sistem güvenliği için geliştirdiğimiz sisteme ait normal durum eğitim süreçleri Şekil 1 özetlenmiştir. Eğitim için OT sistemin saldırılara karşı izole bir ortamda çalışması gerekmektedir. Eğim süreci OT sistem çalışması, verilere erişim ve kayıt, birliktelik analizi ile aktiviteler için önemli olan verilerin tespiti ve aktivite kayıtlarının oluşturulması, aktivite kayıtlarının analizi ile Bayes ağının oluşturulması, OT sistem süreç bilgileri kullanılarak ağ modelinin revizyonu ve son olarak nihai ağ modelinin aktivite kayıtları kullanılarak normal durum senaryoları için eğitimi adımlarından oluşmaktadır. Siber saldırılara açık durumda çalışan OT sistemine ait tehdit analiz safhaları Şekil 2 de yer almaktadır. Aktivite kayıtlarına uygun şekilde veriler OT sisteminden çalışma anında okunur ve aktivite kaydedilir. Oluşturulan her yeni aktivite eğitilmiş Bayes ağında değerlendirilerek normal oluşma olasılığı hesaplanır. Hesaplanan olasılık eşik değerinin altında ise OT sistem yöneticisi uyarılır. Tüm sistemin çalışma mimarisi Şekil 1 ve 2 de özetlenmektedir.

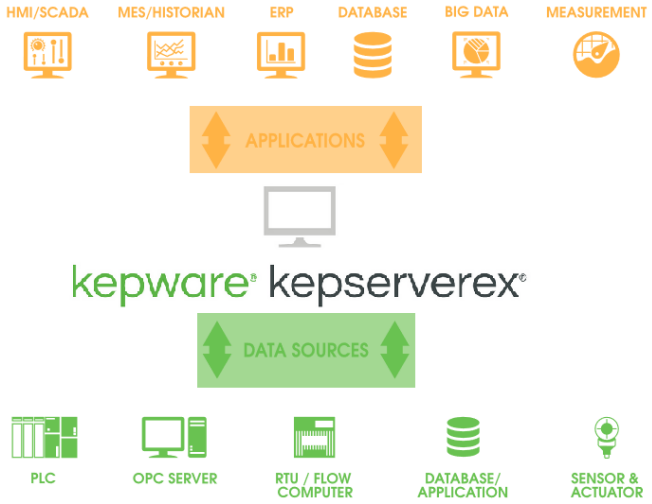


Şekil 1 Sistem Eğitimi

Şekil 2 OT Aktivite Değerlemdirme

2.1. OPC Veri İzleme

OPC, endüstriyel otomasyon alanında ve diğer endüstrilerde güvenli ve güvenilir veri alışverişi için birlikte çalışabilirlik standardıdır. Platformdan bağımsızdır ve birden çok tedarikçinin cihazları arasında kesintisiz bilgi akışını sağlar. OPC Vakfı, bu standardın geliştirilmesi ve sürdürülmesinden sorumludur [21]. OPC standardı, endüstri satıcıları, son kullanıcılar ve yazılım geliştiricileri tarafından geliştirilen bir dizi spesifikasyondur. Bu özellikler, gerçek zamanlı verilere erişim, alarmların ve olayların izlenmesi, geçmiş verilere ve diğer uygulamalara erişim dâhil olmak üzere istemciler ve sunucular ile sunucular ve sunucular arasındaki arabirimi tanımlar [22]. Standart 1996'da ilk kez piyasaya sürüldüğünde, amacı PLC'ye özgü protokolleri (Modbus, Profibus, vb.) standartlaştırılmış bir ara yüze soyutlamaktı. Sonuç olarak, son kullanıcıların, tümü OPC aracılığıyla sorunsuz bir şekilde etkileşime giren, türünün en iyisi ürünleri kullanabilme imkânı tanıyan bir ürün endüstrisi ortaya çıktı. Bugün OPC kısaltması Açık Platform İletişimi anlamına gelmektedir [23]. Bu çalışma kapsamında OT sistemleri ile haberleşme için, fabrika çapında iletişimden tekil IoT cihaz bağlantısına kadar geniş yelpazede ürün bağlantı desteği sağlayan ve API kütüphaneleri ile yazılımsal kontrole olanak tanıyan Kepware OPC [24] çözümü kullanılmıştır. Kepware OPC sunucunun genel mimarisi Şekil 3 de özetlenmektedir.



Şekil 3Kepware OPC Mimari

2.2. Birliktelik Kural Analizi

Birliktelik kural analizi değişkenler arasındaki ilişkileri keşfetmek için kural tabanlı bir makine öğrenimi yöntemidir. Değişkenler arası ilişkileri belirlemeyi amaçlamaktadır [25]. Bunun en yaygın örneği piyasa sepeti analizidir. Sepet analizinde yer alan kayıtlar benzersiz bir tanımlayıcı id ve belirli bir müşteri tarafından satın alınan bir dizi öğeyi içeren işlemlere karşılık gelir. Bu analiz müşterilerinin satın alma davranışları hakkında bilgi edinmek için yaygın kullanılmaktadır. Analizin temel amacı ilişkili ürünlerin tespit edilmesidir. Örneğin soğan ve patates alan müşterilerin bu ürünler ile birlikte ekmek aldığı bir veri setinden beklenen sonuç soğan ve patates ile ekmek satışı arasında güçlü bir ilişki olduğunu yönünde olmasıdır [26]. Birliktelik analizi biyoinformatik, tıbbi teşhis, web madenciliği ve bilimsel veri analizi gibi diğer uygulama alanlarına da uygulanabilir.

I = {i_1, i_2,… i_n, } n adet ikili nitelik içeren öğe kümesi ve D = {t_1, t_2,… t_n, } öğelerin durumlarını içeren işlemler listesi olarak tanımlanmıştır. D'deki her işlemin benzersiz bir işlem kimliği vardır ve I'deki öğelerin bir alt kümesini içerir. Bir kural X=>Y, X,Y⊆ I olarak tanımlanır. Her kural, X ve Y olarak da bilinen iki farklı öğe setinden oluşur; burada X, öncül veya sol taraf ve Y ardışıl veya sağ taraf olarak adlandırılır. Tanımlanan kümelerden birliktelik analiz için birçok algoritma bulunmakla beraber literatürde yaygın kullanılan algoritma Apriori Algoritmasıdır [27]. Algoritma iteratif bir algoritmadır ve veri tabanında sık geçen öğelerin tespit edilmesi için kullanılır. Algoritmada ilk olarak tüm sık geçen öğeler bulunur daha sonra bu öğeler kullanılarak birliktelik kuralları üretilir. Algoritma minimum Destek ve minimum Güven parametre değerlerine göre filtreleme yaparak çalışır. Destek değeri, bir ilişkinin tüm işlemler içinde hangi oranda tekrar ettiğini temsil eder. Güven değeri, X öğesi geçen bir işlemde Y öğesi geçme olasılığını belirtir. Destek ve Güven değerleri aşağıdaki formülasyon ile hesaplanır.

$$Destek(X \rightarrow Y) = \frac{Frekans(X,Y)}{N}$$

$$Güven(X \rightarrow Y) = \frac{Frekans(X,Y)}{Frekans(X)}$$

Algoritmanın işlem basamakları aşağıda yer almaktadır:

- Minimum Destek ve Minimum Güven parametre değerlerini oku.
- Tüm veri setini tara, her bir öğe için tekil, ikişerli, üçerli, vb. olarak grupla, Destek değerlerini hesapla, Minimum Destek parametre değerinden küçük değerlere sahip olan öğeleri çözümden çıkar.
- Kalan d adet öğeyi kullanarak 2^d adet aday Nitelik kümesini oluştur.
- Her aday için Güven değerini hesapla, Güven parametre değerinin üzerinde olan öğeler ile kuralları oluştur.

2.3. Aktivite Tespit Algoritması

OT sistemlerinde birçok giriş-çıkış ve iletişim sinyali bulunmaktadır. Tüm verilerin filtrelenmeden veya ilişkilendirilmeden kullanılması bazı boyuttaki OT sistemler için mümkün olmayacağı gibi ilişkisiz verilerin analizi anormal aktivite tahminin olumsuz yönde etkileyecektir. Verilerin ilişkilendirilmesi ve tanımlanması hedef OT sistem süreçlerine hâkim kullanıcılar tarafından yapılabilir. Bu işlem hem gerektirdiği iş yükü, eğitimli personel hem de insan kaynaklı olası hataların tahminlemeyi direk etkileyeceği sebebi ile tercih edilmemiştir. Ayrıca manuel ilişkilendirme yapılma zorunluluğu beraberinde her OT sistemi için özel tanımlamayı da getirmekte ve sistemin kullanımı zorlaşmaktadır.

Otomatik aktivite analizi için geliştirdiğimiz algoritmamız birliktelik kural analizine dayanmaktadır ve Apriori algoritmasını kullanmaktadır. OPC sunucu tarafından sağlanan tüm ikili OT sistem verileri nitelik olarak kullanılır. OT sisteminden erişilecek olan öğeler kullanıcı tarafından ayarlanmaktadır. I = {i_1, i_2, … i_n, } n adet ikili nitelik içeren öğe kümesini temsil etmek üzere I kümesinde yer alan her i öğesi OT sistemindeki hedef ikili veriyi (vana açık/kapalı, motor çalışıyor/çalışmıyor, sinyal aktif/pasif vb.) göstermektedir.

Algoritmanın yaygın kullanıldığı sepet analizi vb. alanlardan farklı olarak OT sistemlerde olaylar zaman kaymalı şekilde gerçekleşebilmektedir. Yani bir sinyal aktif olup pasife düştükten belirli bir süre sonra bir motor devreye girebilir. Bu durum olayların gerçekleştiği anda bazı öğelerin durumlarının yanlış olmasına ve ilişkilerin kaçırılmasına sebep olur. Bu sonunu çözmek için tanımlanan her öğenin aktiflik gecikme ve pasiflik gecikme tabloları oluşturulur. Tablo her öğenin ne kadar süredir açık, ne kadar süredir kapalı olduğunu gösteren ve i_n_od, i_n_fd olarak adlandırılan ilave değerler içerir. Tabloda yer alan i_n_od, i_n_fd değerlerindeki 0'dan pozitif süreye ve pozitif süreden 0 a geçişler 1 olarak işaretlenerek geçişleri temsil eden ikili veri oluşturulur. Pozitif ve negatif durumları temsil eden ve i_n_od_pt, i_n_od_nt, i_n_fd_pt, i_n_fd_nt öğeri I öğe kümesine eklenir. Örnek öğeye ait geçiş veri hesaplaması Tablo 1 de gösterilmiştir. Tabloda yer alan T sütunu kaydın alındığı zamanı temsil etmektedir.

D = {t_1, t_2, … t_n} öğelerin durumlarını içeren işlemler listesi olmak üzere her işleme benzersiz bir işlem numarası atanır. D kümesinde yer alan her işlem OT sisteminin belirli bir zamanında kayıtlı tüm öğelerin durumu okunarak oluşturulur. Bir işlem kaydının oluşması için koşul bir önceki işlem kaydından farklı olmasıdır. Algoritma OT sistem verilerini OPC üzerinden periyodik olarak okur, okunan işlem kaydını son kaydedilen kayıt ile karşılaştırır ve değişiklik mevcut ise okunan kaydı yeni kayıt olarak saklar aksi durumda kayıt saklanmaz. Siber saldırıdan izole ortamda çalıştırılan OT sistemi

kullanılarak veriler toplanır ve işlem kayıt veri tabanı oluşturulur. Oluşturulan veri tabanı üzerinde Apriori algoritması kullanılarak ilişkili öğeler tespit edilir gerçek zamanlı okuma için hedef öğe listesi oluşturulur. Anomali tespit sistemlerinin temel dezavantajı yüksek yanlış pozitif oranıdır. OT sistemlerinin sahip olduğu öğe sayısı sebebi ile bu oran daha fazla artmaktadır. OT sistemlerinin sahip olduğu avantaj ise yüksek tekrarlı, periyodik iş akışlarıdır. Her kontrol belirli ve kısıtlı şartlar altında çalıştığı için koşulların farklı kombinasyon ile oluşma olasılığı düşüktür. Bu özelliği yanlış pozitif oranlarını düşürmede kullanmak için algoritmamız minimum güven oranı %95 olarak çalıştırılmıştır. OT sistemler üretim süreçlerine bağımlı olarak verilerini kontrol eder bu sebep ile bazı işlem kayıtlarına nadir rastlanabilir. Nadir rastlanan işlem kaydının siber saldırı olma olasılığı ile normal çalışma rutini olma olasılığı aynıdır. İşlem kayıtlarındaki kaybın önlenmesi için minimum destek değeri %5 olarak kullanılmıştır. Aktivite tespit algoritmasının işlem basamakları aşağıda yer almaktadır.

- OPC den veri oku.
- Yeni veriyi bir önceki ile karşılaştır; değişim yok ise bir sonraki okumayı bekle, değişim var ise yeni veri kaydını oluştur.
- Yeni veri için pozitif/negatif geçişleri hesapla.
- Belirlenen süre veya adet okumaya ulaşılmadı ise okumaya devam et, ulaşıldı ise işlem veri tabanını oluştur.
- Tüm veri tabanı için destek değerlerini hesapla, %5 destek değerinden küçük değerlere sahip olan öğeleri sil.
- Tüm öğeler için güven değerini hesapla, %95 güven değerinden küçük olan kayıtları sil.
- Kural kümesini oluştur ve sakla.
- Gerçek zamanlı kontrol için OPC hedef öğe listesi I_{RT} oluştur.

Tablo 1. Pozitif Ve Negatif Geçiş Veri Hesaplama Örneği

i_n	T	i_n_od	i_n_fd	i_n_od_pt	i_n_od_nt	i_n_fd_pt	i_n_fd_nt
0	0	0	0	0	0	0	0
0	3	0	3	0	0	1	0
0	5	0	5	0	0	1	0
1	10	0	10	0	0	1	0
0	20	10	0	1	0	0	1
0	50	0	30	0	1	1	0
0	52	0	32	0	0	1	0
1	57	0	37	0	0	1	0
0	67	10	0	1	0	0	1

2.3. Bayes Ağı

Bayes ağları, verilerden veya uzman görüşlerinden modeller oluşturmak için kullanılan bir Olasılıksal Grafik Modelleme türüdür [28]. Oluşturulan model ile yapılabilecek tahminler, anormallik tespiti, teşhis, belirsizlik altında karar verme gibi çok alanda kullanılabilmektedir. OT sistem için Bayes ağı, bir problem alanı içindeki neden-sonuç ilişkilerinin yönlü çevrimsiz çizgeler ile grafiksel gösterimidir. Düğümler ilgili öğeleri ve yönlü bağlantılarda öğeler arasındaki ilişkiyi temsil eder. Olasılık hesaplamaları için Bayes çıkarımını kullanılır. Bayes ağı, G = (V, E) ile gösterilmiştir. Buradaki V düğümleri yani OT sistemdeki öğeleri ve E kenarları yani öğeler arasındaki ilişkileri temsil eder. Her $v_i \in V$ aynı zamanda birliktelik kural analizi ile oluşturulan öğe listesinin elemanıdır $v_i \in I_{RT.}$ G ağının içerdiği kenar kümesi E Bayes ağı oluşturma safhasında belirlenmektedir. Bayes ağı, bir dizi öğeye (v_1, ..., v_n) ait ortak olasılık yoğunluğunun özel bir temsilidir ve zincir kuralı olarak adlandırılır [29].

$$P(v_1, \dots, v_n) = \prod_{i=1}^{n} P(v_i | \, Pa(v_i))$$

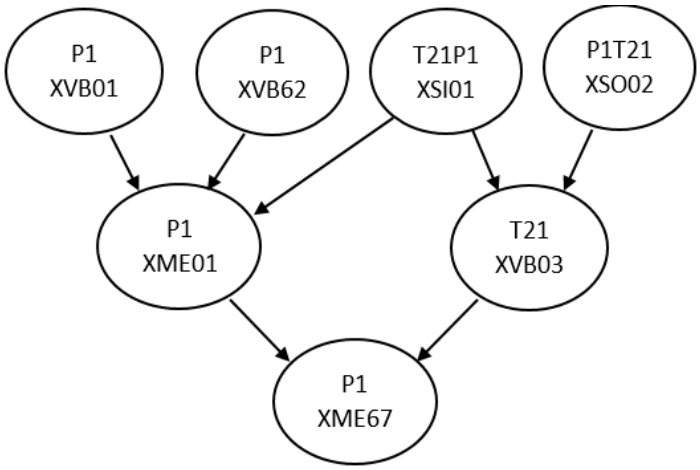
Burada Pa(v_i), v_i düğümünün tüm bağımlılıklarını yani ebeveynlerini temsil eder. Bayes ağını oluşturma adımları aşağıda listelenmiştir:

- Tüm öğelerin listesi belirlenir $I_{RT} = (v_1, \dots, v_n)$.
- Zincir kuralı uygulanır
- Her vi için koşullu bağımsızlıklar değerlendirilerek minimum Pa(vi) oluşturulur
- Belirlenen ebeveyn düğümler ile zincir kuralı tekrar hesaplanır ve Bayes ağı oluşturulur

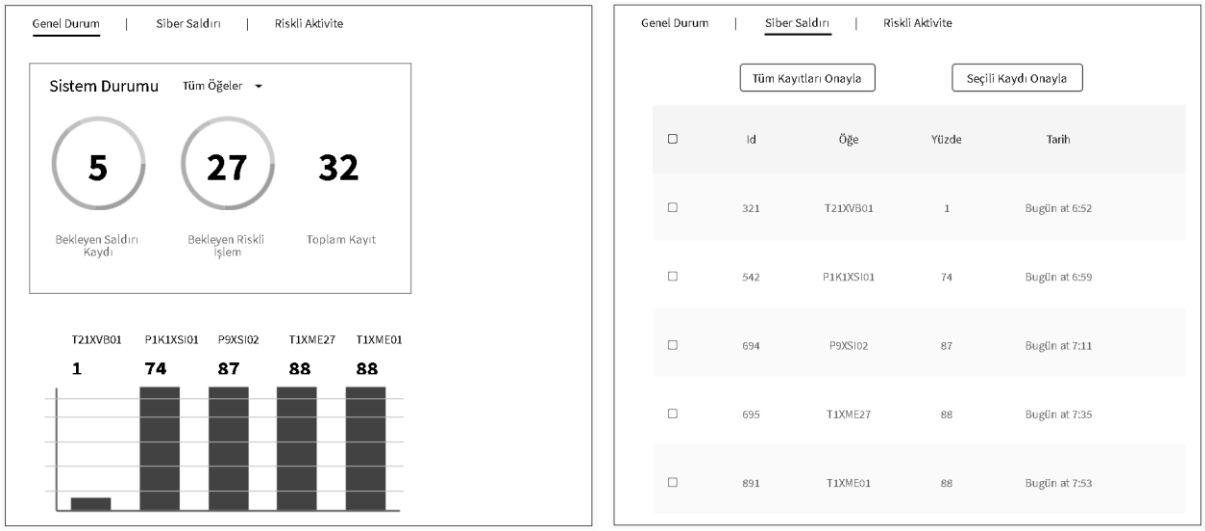
Bayes ağının oluşturulması, değiştirilmesi ve değerlendirilmesi için grafiksel ara birim ve harici program geliştirme araçları ile programlama olanağı sağlayan açık kaynaklı Weka [30] veri madenciliği aracı kullanılmıştır. Aktivite tespit algoritmamız tarafında oluşturulan hedef öğe listesi I_{RT} içerisinde yer alan ve OT sistemde gerçek bir veriyi temsil eden (sistem tarafında üretilen gecikme değerleri hariç) tüm öğeler için Weka Java api'leri kullanılarak Bayes ağları oluşturulmuştur. Ağların oluşturulması aşamasında weka.classifiers.bayes.BayesNet algoritmasında "estimator" parametresi için "SimpleEstimator –A 0,5", "searchAlgorithm" parametresi için "K2 –P 1 –S BAYES" değerleri kullanılmıştır. Bu parametre değerleri Weka tarafından önerilen ön tanımlı değerlerdir. Her öğe için oluşturulan ağaçlar öğe adları ile XML BIF formatında saklanmakta ve Weka ortamı ile açılarak değiştirilebilmektedir. Ağaçlar üzerinde değişiklik işlemi için OT sistem ve ilgili süreç hakkında ön bilgi gereklidir. Bu şekilde oluşturulan ağaçlardan olası bağlantı ve öğe bilgileri silinerek tahminleme performansı iyileştirilmektedir. Oluşturulan Bayes ağlarına müdahale sistem performansını direk etkileyeceği için değişikliği gerçek zamanlı çalışma sırasında oluşan yanlış pozitif değerlerinin incelenmesi ardından yapılmasını öneriyoruz. Bu çalışma kapsamındaki ağ değişikliklerinin tamamı yanlış pozitif değerlerini iyileştirme amacı ile yapılmıştır. Oluşturulan örnek Bayes ağı Şekil 4 de yer almaktadır. Öğe isimleri kullanılan

sisteme özgü olup P1XME67 ürün basma motorunu temsil etmektedir.

Sistemin gerçek zamanlı takibi ve anormal durumların tespiti için Java tabanlı bir uygulama geliştirilmiştir. Uygulama oluşturulan öğe takip listesini OPC sunucu bağlantısı ile periyodik olarak okumaktadır. Okunan değerler ile bir öncekiler karşılaştırılmakta ve sistemde değişiklik olup olamadığı değerlendirilmektedir. Değişiklik olması durumunda ilgili okuma aktivite kaydına dönüştürülmektedir. Aktivite kaydına dönüşün kayıtta yer alan tüm öğeler incelenir ve hangi öğelerin durum değiştirdiği tespit edilir. Durum değiştiren öğelere ait Bayes ağlarına erişilir ve ilgili aktivite kaydı için öğe durumunun oluşma olasılığı hesaplanır. Hesaplanan olasılık değerleri için 3 farklı kategori tanımlanmıştır. Bu kategoriler; güvenilir: %95-%100, riskli: %90-%95 ve saldırı: < %90 değerlerinden oluşur. Program her aktivite kaydını değerlendirerek veri tabanında saklar ve ara biriminde yer alan iki farklı pencerede görüntüler, bunlar saldırı ve riskli uyarı pencereleridir. Oluşan kayıtlar periyodik tarama ile veya canlı olarak incelenir ve yanlış pozitif kayıtlar tespit edilir. Yanlış pozitif kayıtların iyileştirilmesi için iki yol bulunmaktadır. Bunlar yanlış kayıtların eğitim veri setine ilave edilerek mevcut ağların tekrar eğitilmesi ve otomatik oluşturulan ağların revize edilmesidir. İşlemlerin tamamı Weka ortamı kullanılarak yapılabilmekte ve sistem performansı iyileştirilebilmektedir. Geliştirilen uygulamaya ait ekran görüntüsü Şekil 5 de yer almaktadır.



Şekil 4Örnek OT Bayes Ağı



Şekil 5 Geliştirilen Programa Ait Ekran Görüntüleri, Genel Durum Ve Siber Saldırı Sayfaları

Tablo 2. P1XME67 Öğesine Ait Bayes Ağ Weka Performans Sonuçları

Sınıfa Göre Ayrıntılı Doğruluk									Karışıklık Matrisi	
Gerçek Pozitif	**Yanlış Pozitif**	**Kesinlik**	**Duyarlılık**	**F-Ölçümü**	**MCC**	**ROC Alan**	**PRC Alan**	**Sınıf**	**A=1**	**B=0**
0,888	0,083	0,944	0,888	0,915	0,793	0,970	0,983	1	237	30
0,917	0,112	0,837	0,917	0,875	0,793	0,970	0,952	0	14	154
0,899	0,095	0,903	0,899	0,900	0,793	0,970	0,971			

3. Araştırma Sonuçları ve Tartışma

Geliştirilen sistemin performans testleri için OT sistem 1 gün boyunca çalıştırılıp kayıtlar toplanarak Bayes ağ eğitimi gerçekleştirilmiştir. Şekil 4 de yer alan "P1XME67" öğesine ait eğitim performans sonuçları ve karmaşıklık matrisini gösteren Weka sonuç tablosu Tablo 2 de yer almaktadır.

Sistemin başlangıç koşullarındaki genel eğitim performansının hesaplanması için OT sistemde yer alan tüm öğelerin performans sonuçlarının ortalaması hesaplanmıştır. Sistemin ortalama eğitim performans sonuçları Tablo 3 de sunulmuştur. Performans değerlerinden görüleceği üzere tüm öğeler göz önüne alındığında sistemin yanlış pozitif oranları artış göstermektedir.

Tablo 3 Sistemin İlk Eğitim Sonrası Ortalama Eğitim Performansı

Duyarlılık	Özgüllük	Kesinlik	F-Ölçümü
0,9091	0,7778	0,7143	0,8000

Sistem eğitiminde OT sistemde manuel operatör işlemleri yapılmamıştır. Bu durum OT sisteme manuel müdahalelerin anormal sayılmasına ve saldırı olarak işaretlenmesini sağlayacaktır. Bu şekilde normal saldırılara ilave olarak operatörlerin hatalı veya kasıtlı olarak yaptıkları işlemlerde tespit edilebilir. Sistem, OT sisteminde yer alan tüm öğelerin işlevsel hareketlerini öğrenmektedir ve işlevsel davranışlar OT sistem kontrol programları tarafından düzenlenmektedir. Bu sebep ile OT sistem üzerinden yapılacak her türlü program değişikliğinin oluşturacağı farklı öğe davranışları da sistem tarafından anormal aktivite olarak algılanacaktır. Bu yaklaşım sayesinde OT sisteme program kodu düzeyinde yapılacak saldırılarda tespit edilebilmektedir. Bu durumun dezavantaja dönüşmemesi için OT sistemde yapılan kontrollü kod değişimlerinin ardından sistemin ilgili ağları tekrar eğitilmelidir. Sistemin tespit performansının ölçülmesi için OT sisteme 3 farklı saldırı düzenlenmiştir. Bunlar yetkisiz manuel müdahale, yetkisiz kod değişimi ve OPC ağ trafiği üzerinden modbus haberleşme kanalına yapılan paket değişim saldırılarıdır. Her saldırı türünde 50 farklı öğenin durumuna müdahale edilmiş olup tüm saldırılar sistem tarafından "siber saldırı" kategorisinde kayıt altına alınmıştır. Sistemin yanlış pozitif oranlarının düşürülmesi için OT sistem ilave 1 gün daha çalıştırılarak veri seti genişletilmiş ve Bayes ağları yeni veri seti ile eğitime tabi tutulmuştur. Yapılan çalışma veri setindeki artışın sistemin ortalama performansına pozitif yansıdığı, OT süreç uzmanları tarafından yapılacak sistem eğitimleri ile sistemin yanlış doğru oranının azaltılabileceği gösterilmiştir. İkinci eğitim sonrası sistemin ortalama eğitim performans sonuçları Tablo 4 de sunulmuştur.

Tablo 4. Sistemin İkinci Eğitim Sonrası Ortalama Eğitim Performansı

Duyarlılık	Özgüllük	Kesinlik	F-Ölçümü
0.8750	0.8182	0.7778	0.8235

4. Sonuç

OT sistemlerinin mevcut durumu ve siber güvenlik risklerinin göz önüne alındığı bu çalışmada OT sistemleri için saldırı tespit ve uyarı sistemi geliştirilmiştir. Geliştirilen sistem OT sisteme bağlanarak gerekli verileri okumak için OPC sunucu bağlantı modülü içermektedir. Bu sayede OT sistemde kullanılan teknolojik yazılım ve donanım bağımsız olarak tüm OT sistemlere bağlanabilmektedir. OT sistemlerde meydana gelen işlemlerin algılanması ve analiz edilebilmesi için birliktelik analizi tabanlı aktivite kayıt oluşturma algoritması geliştirilmiş. Geliştirilen bu algoritma ile OT süreç bilgisi olmadan tüm aktivitelerin yorumlanabilmesine olanak sağlanmıştır. Süreç bilgileri kullanılarak algoritmanın okuma yapacağı öğeler değiştirilerek iyileştirilebilmektedir. Oluşan aktivite bilgilerinin analizi için Bayes ağ tabanlı bir öğrenme sistemi tasarlanmıştır. Bu sistem sayesinde elde edilen kayıtlardan Bayes ağları oluşturulmakta, sistem çalışma anında oluşan tüm aktiviteler değerlendirilerek olasılıklarına göre "Güvenli", "Riskli" veya "Siber Saldırı" olarak gruplandırılmakta ve OT yetkililerine sunulmaktadır. Önerilen sistem her tür OT aktivitesine adapte olabilecek yapıdadır. Entegrasyon için OT müdahalesi gerektirmemekte, OT verilerine erişim sağlayacak bir OPC sunucu yada PLC veya SCADA sistem verilerine okuma erişimi yeterli olmaktadır. Sistem mimari yapısı gereği sadece siber saldırı kategorisindeki işlemleri tespit etmekte kalmayıp sistemde yetkili kişilerin hatalı yâda kasıtlı müdahale veya program değişimlerini de algılayabilmektedir. Bu avantajın oluşturduğu dezavantajda saldırı olmayan ve normal koşullar altında gerçekleşen müdahale ve program değişikliklerinin eğitim ile sisteme öğretilme gereksinimidir.

Deneysel çalışmalarımız OT sisteme yapılan her türlü saldırının tespit edilebildiğini göstermektedir. Yüksek doğru pozitif oranlarına karşın tüm anormal durum tespit sistemlerinde olduğu gibi sistemimizde yüksek yanlış pozitif oranlarına sahiptir. Yapılan deneysel çalışmalar eğitim sürecinin arttırılmasının ve OT süreç bilgileri ile toplanan veri ve oluşan Bayes ağaçlarına yapılan müdahalelerin sistem performansını arttırarak yanlış pozitif oranlarını düşürdüğünü ispatlamaktadır.

Her tür bilgisayar sistemi siber saldırılara karşı savunmasızdır. OT sistemler icra ettikleri fonksiyon ve çalışma şartları gereği günümüzde siber saldırılara karşı en savunmasız sistemler haline dönüşmektedir. Bu çalışma ile mevcut OT sistemlerine minimum değişiklik ile eklenebilen ve her tür saldırıya karşı önleyici tedbir sağlayan bir sistem önerilmiştir. Yapılan deneysel çalışmalar ile tasarlanan sistemin her tür OT sisteme uygulanabilecek yapıda olduğu ve saldırı tespitinde başarılı sonuçlar elde ettiği görülmüştür. Bu çalışma OT davranışlarının modellenerek öğrenildiği ve anormal davranışların tespit edilerek siber saldırıların tespit edildiği ilk çalışmadır. Birliktelik Analizi ve Bayes ağları ile elde edilen başarılı sonuçlar, bu yöntemlerin siber güvenlik alanına uygunluğunu göstermesi sebebi ile yapmış olduğumuz çalışma gelecekte bu alanda yapılacak araştırmalara öncülük edebilir.

Kaynakça

1. Mukkamala, S., Sung, A., & Abraham, A. (2005). Cyber security challenges: Designing efficient intrusion detection systems and antivirus tools. Vemuri, V. Rao, Enhancing Computer Security with Smart Technology.(Auerbach, 2006), 125-163.

2. Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2013). Network anomaly detection: methods, systems and tools. Ieee communications surveys & tutorials, 16(1), 303-336.

3. Stouffer, K., Falco, J., & Scarfone, K. (2011). Guide to industrial control systems (ICS) security. NIST special publication, 800(82), 16-16.

4. Framework, S. (2010). Policy Statement on Improving the Resilience of Critical Infrastructure to Disruption from Natural Hazards. London: Cabinet Office.

5. Henrie, M. (2013). Cyber security risk management in the SCADA critical infrastructure environment. Engineering Management Journal, 25(2), 38-45.

6. Guan, J., Graham, J. H., & Hieb, J. L. (2011, July). A digraph model for risk identification and mangement in SCADA systems. In Proceedings of 2011 IEEE International Conference on Intelligence and Security Informatics (pp. 150-155). IEEE.

7. Patel, S., Tantalean, R., Ralston, P., & Graham, J. (2005). Supervisory control and data acquisition remote terminal unit testbed. Intelligent Systems Research Laboratory technical report TR-ISRL-05-01, Department of Computer Engineering and Computer Science. Louisville, Kentucky: University of Louisville, 24, 26.

8. Yan, J., Liu, C. C., & Govindarasu, M. (2011, March). Cyber intrusion of wind farm SCADA system and its impact analysis. In 2011 IEEE/PES Power Systems Conference and Exposition (pp. 1-6). IEEE.

9. Ericsson, G. N. (2007). Toward a framework for managing information security for an electric power utility—CIGRÉ experiences. IEEE transactions on power delivery, 22(3), 1461-1469.

10. Amin, M. (2002). Security challenges for the electricity infrastructure. Computer, 35(4), supl8-supl10.

11. Schneider, K., Liu, C. C., & Paul, J. P. (2006). Assessment of interactions between power and telecommunications infrastructures. IEEE Transactions on Power Systems, 21(3), 1123-1130.

12. Alves, T., & Morris, T. (2018). OpenPLC: An IEC 61,131–3 compliant open source industrial controller for cyber security research. Computers & Security, 78, 364-379.

13. Davis, C. M., Tate, J. E., Okhravi, H., Grier, C., Overbye, T. J., & Nicol, D. (2006, September). SCADA cyber security testbed development. In 2006 38th North American Power Symposium (pp. 483-488). IEEE.

14. Ericsson, G. N. (2010). Cyber security and power system communication—essential parts of a smart grid infrastructure. IEEE Transactions on Power Delivery, 25(3), 1501-1507.

15. Seijo Simó, M., López López, G., & Moreno Novella, J. I. (2017). Cybersecurity vulnerability analysis of the plc prime standard. Security and Communication Networks, 2017.

16. Byres, E., & Lowe, J. (2004, October). The myths and facts behind cyber security risks for industrial control systems. In Proceedings of the VDE Kongress (Vol. 116, pp. 213-218).

17. Xie, P., Li, J. H., Ou, X., Liu, P., & Levy, R. (2010, June). Using Bayesian networks for cyber security analysis. In 2010 IEEE/IFIP International Conference on Dependable Systems & Networks (DSN) (pp. 211-220). IEEE.

18. Coluccia, A., D'Alconzo, A., & Ricciato, F. (2013). Distribution-based anomaly detection via generalized likelihood ratio test: A general maximum entropy approach. Computer Networks, 57(17), 3446-3462.

19. Fronza, I., Sillitti, A., Succi, G., Terho, M., & Vlasenko, J. (2013). Failure prediction based on log files using random indexing and support vector machines. Journal of Systems and Software, 86(1), 2-11.

20. Vickers, N. J. (2017). Animal communication: when i'm calling you, will you answer too?. Current biology, 27(14), R713-R715.

21. Mahnke, W., Leitner, S. H., & Damm, M. (2009). OPC unified architecture. Springer Science & Business Media.

22. Lieping, Z., Aiqun, Z., & Yunsheng, Z. (2007, July). On remote real-time communication between MATLAB and PLC based on OPC technology. In 2007 Chinese Control Conference (pp. 545-548). IEEE.

23. Zheng, L., & Nakagawa, H. (2002, August). OPC (OLE for process control) specification and its developments. In Proceedings of the 41st SICE Annual Conference. SICE 2002. (Vol. 2, pp. 917-920). IEEE.

24. Resnick, C. (2012). Kepware Communication Solutions Help Optimize OPC Connectivity. ARC View.

25. Piatetsky-Shapiro, G. (1991). Discovery, analysis, and presentation of strong rules. Knowledge discovery in databases, 229-238.

26. Agrawal, R., Imieliński, T., & Swami, A. (1993, June). Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on Management of data (pp. 207-216).

27. Agrawal, R., & Srikant, R. (1994, September). Fast algorithms for mining association rules. In Proc. 20th int. conf. very large data bases, VLDB (Vol. 1215, pp. 487-499).

28. Guo, Y., Bai, G., & Hu, Y. (2012, December). Using bayes network for prediction of type-2 diabetes. In 2012 International Conference for Internet Technology and Secured Transactions (pp. 471-472). IEEE.

29. He, J., Bai, S., & Wang, X. (2017). An unobtrusive fall detection and alerting system based on Kalman filter and Bayes network classifier. Sensors, 17(6), 1393.

30. Eibe, F., Hall, M. A., & Witten, I. H. (2016). The WEKA workbench. Online appendix for data mining: practical machine learning tools and techniques. In Morgan Kaufmann.