EŞLER ARASI AĞLARDA MAKİNE ÖĞRENİMİ DESTEKLİ TUTARLILIK TEMELLİ GÜVEN YÖNETİMİ

MACHINE LEARNING AIDED CONSISTENCY BASED TRUST MANAGEMENT IN P2P NETWORKS

YASİN ŞAHİN

DOÇ. DR AHMET BURAK CAN

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Bilgisayar Mühendisliği Anabilim Dalı için Öngördüğü

DOKTORA TEZİ olarak hazırlanmıştır.

2022

Desteğini her zaman hissettiğim,

Bu eserin çıkmasında en önemli ilham kaynağım

Biricik eşim Ceren’e

ÖZET

EŞLER ARASI AĞLARDA MAKİNE ÖĞRENİMİ DESTEKLİ TUTARLILIK TEMELLİ GÜVEN YÖNETİMİ

Yasin ŞAHİN

Doktora, Bilgisayar Mühendisliği Bölümü

Tez Danışmanı: Doç. Dr. Ahmet Burak CAN

Ağustos 2022, 88 sayfa

Dosya paylaşımı, dağıtımı gibi uygulamalarla ilk ortaya çıkan eşler arası (P2P) ağlar, blok zinciri, kripto para birimleri, NFT'ler, sanal evren platformları gibi yeni kullanım alanlarında popülerliğini devam ettirmektedir. Eşler arası ağlar, merkezi olmayan yapıları ile sonsuz ölçeklenebilirlik gibi faydaların yanında, tutarlı, sağlam ve güvenilir bir şekilde hizmet sürekliliğini sağlayacak birtakım kurallara ve mekanizmalara ihtiyaç duyarlar.

Eşler arası ağlar, merkezi olmayan doğası gereği, hizmet sağlayıcıların ve hizmet alıcıların manipülasyonuna açıktır. Blok zinciri gibi güvenilir şifreleme algoritmaları ile bir kaydın değişmeyeceğini garanti eden sistemler için bile, söz konusu kaydın orijinal hali ile doğruluğunu teyit eden ayrı bir güven mekanizmasına gereksinim vardır. Bu nedenle güven yönetimi, uzun süredir çalışılmasına rağmen, eşler arası ağlarda en önemli araştırma alanlarından biri olmaya devam etmektedir.

Bu tezde, eşler arası ağın kötü niyetli davranışlardan korunmasını ve güvenilir bir ortamın sağlanması için ihtiyaç duyulan güven modelleri üzerine çalışılmıştır. Eksiksiz bir model elde etmek için hem hizmet saldırıları hem de geri bildirim saldırıları kapsamlı şekilde araştırılmıştır. İki aşamada gerçekleştirilen çalışmanın birinci aşamasında tutarlılık temelli istatistiksel bir model önerilmiştir. Bu model ile özellikle hizmet saldırılarını önlemede çok iyi başarım elde edilmiş olunsa da, bazı geri bildirim saldırılarını tespit etmekte aynı düzeyde bir başarım elde edilememiştir. Çalışmanın ikinci aşamasında ise farklı makine öğrenmesi yöntemlerini uygulayarak daha geniş saldırı senaryoları için güven modeli geliştirilmiştir. Bu modelde, farklı gruplarda özniteliklerden oluşan bir öznitelik kümesi üzerinde çalışılmıştır. Elde ettiğimiz sonuçlar, tutarlılık temelli modelimizi filtreleme amacıyla kullandıktan sonra, makine öğrenmesi modelimizi uyguladığımızda daha güvenilir bir eşler arası ağ elde edebileceğimizi gösterdi. Ayrıca modelin tutarlılığı özendiren doğasının bir sonucu olarak, güvenilir eşlerin daha yüksek hizmet kalitesi elde ettiği ve güvenilir eşler arasında hizmetlerin genelde daha adil bir şekilde dağıtıldığı gözlemlenmiştir.

Anahtar Kelimeler: Eşler arası ağlar, Güven Yönetimi, Makine Öğrenmesi, Tutarlılık Temelli İstatistiksel Model

ABSTRACT

MACHINE LEARNING AIDED CONSISTENCY BASED TRUST MANAGEMENT IN P2P NETWORKS

Yasin SAHIN

Doctor of Philosophy, Department of Computer Engineering

Supervisor: Assoc. Prof. Dr. Ahmet Burak CAN

August 2022, 88 pages

Peer-to-peer (P2P) networks first emerged with file sharing and distribution applications. P2P networks continue their popularity with new usage areas such as cryptocurrencies, NFTs, and metaverse platforms. While peer-to-peer networks provide benefits such as infinite scalability with their decentralized structure, they need some rules and mechanisms to ensure service continuity in a consistent, robust and reliable manner.

Due to their decentralized nature, P2P networks are vulnerable to manipulation by service providers and service receivers. Even for the systems like blockchain, which guarantee unchangeability of a record by using reliable encryption algorithms, there is a need for a separate trust mechanism that confirms the authenticity of the record in its original state. Therefore, trust management remains one of the most important research areas in P2P networks, although it has been studied for a long time.

In this thesis, the trust models, which are needed to protect the peer-to-peer network from malicious behavior and to provide a reliable environment have been studied. Both service attacks and feedback attacks have been extensively investigated to obtain a complete model. In the first stage of the study, which was carried out in two stages, a statistical model based on consistency is proposed. Although this model has achieved very good performance in preventing service attacks, it cannot show the same level of performance in detecting some feedback attacks. In the second stage of the study, a trust model is proposed for wider attack scenarios by applying different machine learning methods. In this model, a feature set consisting of different groups of features was studied. Our results showed that we can achieve a more reliable peer-to-peer network when we apply machine learning after using our consistency-based model for filtering. It has also been observed that, as a result of the consistency-promoting nature of the model, trusted peers got better quality of service (QoS) and services were distributed among trusted peers more fairly.

Keywords: Peer-to-peer networks, Trust management, Machine Learning, Consistency based statistical model

TEŞEKKÜR

Tez süresince bana destek olan:

Danışmanım Doç. Dr. Ahmet Burak Can'a,

Doktora Tez İzleme Komitesinde yer alan Prof. Dr. Ali Aydın Selçuk, Doç. Dr. Murat Aydos hocalarıma,

Her türlü kolaylığı gösteren değerli Hacettepe Üniversitesi Bilgisayar Mühendisliği bölüm personeline,

Eğitim hayatım boyunca desteklerini hissettiğim tüm öğretmenlerime, hocalarıma, değerli büyüklerime ve bölümde görev aldığım süreçte birlikte çalıştığım mesai arkadaşlarıma,

İlk öğretimden yüksek öğrenime her zaman bana destek olan, varlıkları ile güç bulduğum, yıllar içinde büyüyüp güzelleşen aileme,

Sonsuz teşekkürlerimi sunuyorum.

Yasin Şahin

Temmuz 2022, Ankara

İÇİNDEKİLER

ÇİZELGELER DİZİNİ

ŞEKİLLER DİZİNİ

TERİMLER VE KISALTMALAR

Terimler

Accuracy	Doğruluk
Archiver	Arşivi
Attack Rate	Saldırı Oranı
Attribute	Özellik
Blockchain	Blok zinciri
Bootstrap	Önyükleme
Classification	Sınıflandırma
Client	İstemci
Cluster	Küme
Collaborative	İşbirlikçi
Consistency	Tutarlılık
Credibility	İtibar
Error Rate	Hata Oranı
Feature	Öznitelik
Feedback	Geri Bildirim
Free Rider	Kaytaran (P2P ağdan yalnızca hizmet alan, hizmet vermeyen eşler)
Fuzzy Logic	Bulanık Mantık
Interaction	Etkileşim
Machine Learning	Makine Öğrenmesi
Malicious	Kötücül
Peer	Eş

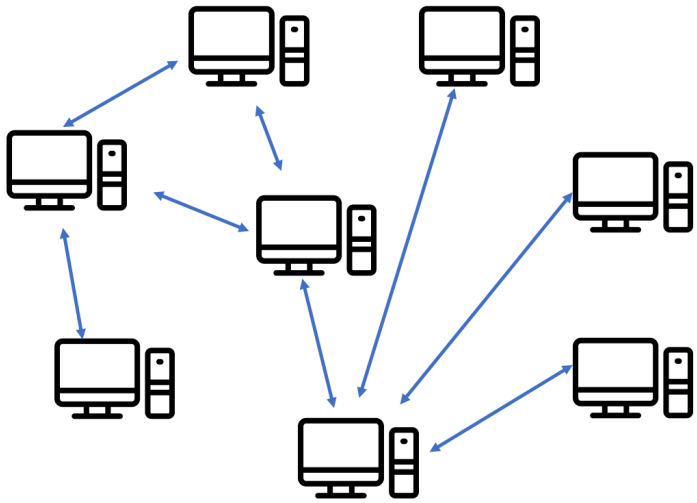
Reputation	İtibar
Satisfaction	Memnuniyet
Sensor	Algılayıcı
Service	Hizmet
Similarity	Benzerlik
Threshold	Eşik
Trust	Güven
Tuple	Demet
Whitewashing	Aklama

Kısaltmalar

CPU	Central Process Unit
DHT	Distributed Hash Table
DS	Dempster-Shafer Evidence Theory
HMM	Hidden Markov Model
IOT	Internet of Things
KNN	K-Nearest Neighbour
ML	Machine Learning
MLE	Maximum Likelihood Estimation
MLP	Multi Layer Perceptron
NBC	Naive Bayes Classifier
NFT	Non-fungable Token
P2P	Peer to Peer
QoS	Quality of Service
SVM	Support Vector Machine
WEKA	Waikato Environment for Knowledge Analysis

1. GİRİŞ

Çevrimiçi istemciler her geçen gün artarken, talep ettiği kaynak çeşitliliği ve miktarı da katlanarak yükselmektedir. Bu artışın hem ağ trafiği ihtiyacı oluşturduğu hem de donanım gereksinimleri merkezi bir noktada toplayamayacak kadar genişlettiği ortadadır. Eşler arası ağlar, ağ trafiğini ve donanım gereksinimlerini istemcilere dağıtan, istemcileri birer sunucu haline getiren ağlardır.

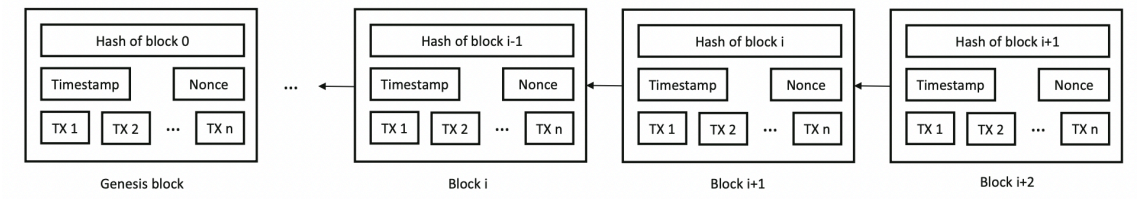


Şekil 1.1 Temel bir eşler arası ağ topolojisi

Eşler arası ağlar ve güven yönetimi farklı kurgularda hizmeti sağlayan ile hizmeti tüketeni bir araya getiren, derecelendirmelerin olduğu sistemlerde yaygın olarak kullanılmaktadır. Örneğin alışveriş yaptığınızda satıcıyla ilgili verdiğimiz geri bildirim ya da bu geri bildirimlerden elde edilen skor, Eşler arası ağlardaki güven yönetimi için iyi bir benzetimdir. Bu anlatımda herkesin dürüst davrandığı ve elde edilen skorun şüpheye yer bırakmadığı düşünülebilir. Ancak bir hizmetin veya ona karşılık verilen geribildirimin gerçekten kötücül olup olmadığı yalnızca hizmeti veren ile alanın arasındaki bir bilgidir. Bu yüzden her türlü manipülasyona açıktır. Şüphesiz ki bu sistemleri sürdüren şirketler farklı isimlerle ekipler kurarak platformun güvenilir işlemesini sağlayan algoritmalar koşturmakta, durumları değerlendirerek skorlara insan değerlendirmesi sonucu elle müdahale edebilmektedir. Ne var ki bu örnekte, sistemin tamamen merkeziyetsiz olmadığı, güvenilir bir otoritenin var olduğu bir senaryodan söz ediyoruz. Söz konusu

otoriteyi ve merkezi veri deposunu sistemden çıkardığımızda, kendi kendini yürütmesi beklenen sistemin güvenli bir ortam sağlamasının zorluğunun artacağından söz edebiliriz.

Eşler arası ağları, alışveriş sistemleri dışında CPU paylaşımının yapıldığı mobil ağlarda, içerik paylaşımı yapılan ortamlarda ya da son yıllarda hayatımıza çok hızlı giriş yapan blok zinciri teknolojilerinde yaygın olarak görebiliyoruz. Blok zinciri tamamen merkeziyetsiz bir ağ mimarisi üzerine kurgulanmış eşler arası ağ örneğidir. Blok zinciri teknolojisi verinin değişmezliğini ve kalıcılığını sağlayan bir çözümdür. Şekil 1.2'de gösterilen blok zincir yapısında, sıralı olarak birbirini işaret eden blokların değişmezliği matematiksel olarak çözülmesi güç algoritmalar ile sağlanır [1]. Blok zinciri çözümleri kargo takibi, perakende zinciri, para transferi, kripto para varlıkları, NFT'ler vb. pek çok alanda başarı ile uygulanmaktadır. Blok zinciri teknolojilerinin sistem gereksinimine göre sürekli büyüme sorunu dışında önemli bir güven yönetimi sorunu da vardır. Zira bu zincirler içerdikleri verinin değişmezliğini güvence altına almakla birlikte, verinin kendisinin doğruluğunu garanti etmezler. Yani, okuma sırasında, kaydın değiştirilmemiş olduğunu garanti ederken, yazma sırasında, kayıt altına alınması talep edilen işlemin doğruluğu için ayrı sistemlere ihtiyaç duyulabilir.



Şekil 1.2 Örnek bir blok zinciri yapısı [2]

Eşler arası ağların uygulandığı kurgulardan bazıları şifreleme ile herhangi bir güven ilişkisine dayanmazken pek çoğu hem hizmet hem deneyim paylaşımında güven yönetimine ihtiyaç duyarlar. Hiç tanımadığımız birine koşulsuz güvenmek mümkün olmadığı gibi eşler arası ağlarda da istemciler arası kaynak paylaşımının yapılabilmesi için karşılıklı onay sürecine, bir başka deyişle güven yönetim mekanizmalarına ihtiyaç duyulur. Bir istemcinin bir hizmeti onaylamasındaki genel işleyiş; geçmiş deneyimlerin toplanması, istemci eşin kendi deneyimi ile ilişkilendirilmesi ve hizmet sağlayacak olan eşe güvenip güvenmeyeceğine karar vermesi şeklinde gerçekleşmektedir.

Güven, bulanıklığın çok yüksek olduğu, fen bilimlerinden ziyade sosyal bilimlerin incelediği bir kavramdır. Ne var ki henüz üzerinde uzlaşılmış tek bir tanıma sahip de değildir. Halis ve Şenkal, güveni;

> "*Güven En kısa biçimde, "olumlu beklentilere sahip olarak bir başkasının etkilerine açık olma niyetini kapsayan psikolojik bir durum" ya da "bir başkasının tavırlarına yönelik sahip olunan olumlu beklentilerden emin olma" ve kişisel risk bağlamında "karşı tarafın yapıcı bir davranış sergileyeceğine dair inanç ya da beklenti" olarak ifade edilebilir."*

olarak tanımlamaktadır [3]. Güveni iki ayrı biçimde gruplayabiliriz; kişiselleştirilmiş güven ve sosyal/genelleştirilmiş güven.

- *Kişiselleştirilmiş Güven:* Bir kişinin aile bireyleri, dostları, iş arkadaşları ve komşusu gibi tanıdığı insanlara karşı duyduğu güven duygusudur [4]. Güven yönetiminin eşler arası ağlarda araştırıldığı çalışmalarda zaman zaman yer verilen yerel güven değeri yaklaşımlarını kişiselleştirilmiş güven olarak gruplandırabiliriz.
- *Sosyal/Genelleştirilmiş Güven:* Bir kişinin toplumda tanımadığı insanlara karşı duyduğu güven duygusudur [4]. Eşler arası ağlarda güven yönetiminde genel güven değeri, itibar gibi farklı şekillerde uygulandığını gördüğümüz yaklaşımları sosyal/genelleştirilmiş güven olarak gruplandırabiliriz.

Eşler arası ağların merkezi olmayan doğası, güven ilişkilerinin eşler arasında kurulmasını zorlaştıran bir etmendir. Bu zorluk bu alandaki çalışmaların odak noktasını oluşturur. Genelde eşler arası ağlarda bire bir ilişkilerden elde edilen deneyim ile diğer eşlerin deneyimlerini geri bildirim ile sunması üzerinden güven kararı veriliyor. Güven kararının verilmesi aşamasında kullanılan geri bildirimlerin yanıltıcı bilgi içermesi olasılığından dolayı matematiksel formüllerle kolayca hesaplanabilecek güven ölçütleri oluşturmak çözülmesi zor bir problemdir.

Bu tez kapsamında yaptığımız çalışmaların birinci aşamasında davranışların ve etkileşimlerin tutarlılığına dayalı bir güven mekanizması oluşturarak hem hizmet sağlayan eşlerin saldırılarını hem de hizmet alan eşlerin değerlendirme yaparken gerçekleştirdiği geri bildirim saldırılarını tespit etmeye çalıştık. Bu birinci aşamada sunduğumuz çözümde, hizmeti almak isteyen ve sağlayabilecek eşler karşılıklı olarak birbirlerine güvenmedikçe etkileşime başlamıyor, bunu da karşı eşin ne kadar tutarlı/güvenilir davranış sergilediğini istatistiksel bir model ile hesaplayarak yapıyoruz. Çalışmamızda pek çok saldırı türünde başarı ile eşleri saldırılardan korumayı başarmış olsak da bazı saldırı türlerinde nispeten açık noktalar olduğunu tespit ettik. İkinci aşamada ise makine öğrenimi yöntemlerinden faydalanarak, hem modelimizdeki açık noktaları olabildiğince kapatmayı hem de başarımın yüksek olduğu saldırı türlerinde daha yüksek başarım sağlamayı hedefledik. Geçmiş etkileşimlerin oluşturduğu tarihi verilere dayanan geleneksel yöntemlerin güvenilirliği ile geçmiş verilerin olmadığı başlangıç sürecinin ele alınmasını destekleyen makine öğrenmesi yöntemlerini birleştirdiğimiz bu çalışmamızda başarılı bir birliktelik elde ettiğimizi söyleyebiliriz. Yaptığımız çalışmalar sırasında gözlemlediğimiz, sıkıntısını çok fazla hissettiğimiz ve çalışmalarımızda ciddi zamanı ayırmak zorunda kaldığımız veri kümesi elde etme ve karşılaştırmak istediğimiz modellerin karşılaştırmalarda kullanılamaması sorunlarına çözüm getirebilmek, bu eksiğin giderilmesine katkı sağlayabilmek adına; Peersim [5] benzetim ortamı üzerinde geliştirilen çalışmanın kaynak kodları açık olarak sunulacaktır. Ayrıca, çalışmanın kodunda mümkün olan her şey yapılandırılabilir, açılıp kapatılabilir şekilde sağlanacaktır. Böylece, gelecekte yapılacak çalışmalar için karşılaştırma ortamı olarak kullanılabilecek bir ortam araştırmacılara sağlanmış olacaktır.

Bu tezin 2. bölümünde bu alandaki farklı yaklaşımlar tartışılacak ve bu çalışmada faydalanılan yöntemler ile araçlar anlatılacak, 3. bölümde tutarlılık temelli bir istatiksel model açıklanacak ve bu modelle ilgili yapılan benzetim temelli analiz çalışmaları verilecektir. 4. bölümde tutarlılık temelli istatistiksel model üzerine geliştirdiğimiz makine öğrenmesi destekli modeli anlatarak 5. bölümdeki deneylerimize ve sonuçlarına geçeceğiz. Son bölümde çalışmamızdan elde ettiğimiz son sözümüzü söyleyeceğiz

2. GENEL BİLGİLER

Eşler arası ağlarda güven kavramı aynı zamanda ağın hizmet kalitesini de belirleyen belki en önemli bileşenidir. Güven yönetimi, eşler arası ağlar ve tamamen ya da kısmen eşler arası altyapı üzerine kurgulandığını söyleyebileceğimiz çoklu-etmen sistemleri, sosyal ağlar, e-ticaret sistemleri gibi alanlarda yaygın olarak sistem kalitesini ve güvenliğini sağlamak için çalışılan bir konudur.

Geleneksel yaklaşımlarda güven yönetimini, ağın yapısına göre iki ana başlığa ayrılır. Yapısal eşler arası ağlar ve yapısal olmayan eşler arası ağlar bu iki ana başlığı oluştururken, yapısal olmayan ağların küçük değişikliklerle kısmi yapısal ağlar olarak konumlandırıldığı çözümler de mevcuttur. Ancak ilerleyen zamanlarda çalışmalar ağın yapısından ziyade güven yönetimine getirdikleri modern çözüm yaklaşımlarına göre sınıflandırılmaya başlanmıştır.

2.1. Dağıtılmış Anahtar Çizelgesi

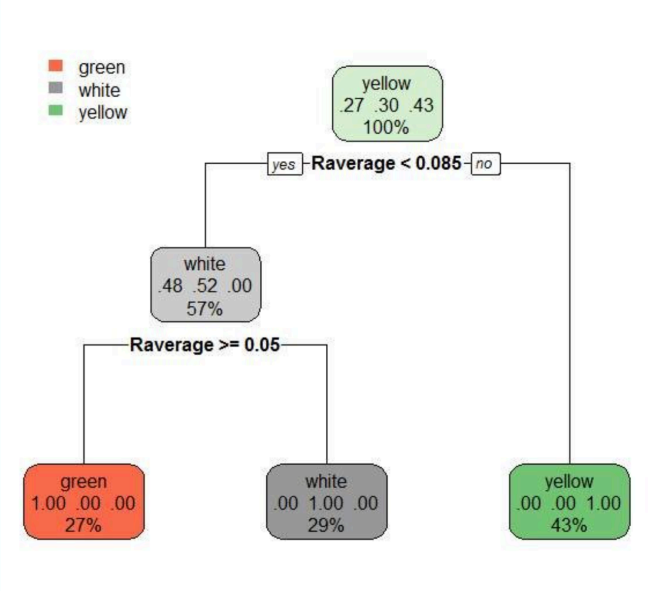
Dağıtılmış Anahtar Çizelgesi (*DHT – Distributed Hash Table*), verilerin eşler arası ağlar üzerinde tutulmasını sağlayan bir tür veri yönetimi çözümüdür. Veriler anahtar-değer (*key-value/data*) ikilileri olarak tutulur ve her anahtardan sorumlu bir eş vardır. Bu sorumluluk, ağın ne kadar güvenilir ve tutarlı olmasının arzu edilmesine bağlı olarak değişen sayıda yedeklenerek eşlere dağıtılır. Yani bir eş bir anahtar grubundan asil/efendi olarak sorumluyken, her bir anahtarın sorumluluğu yedek üyeler olarak değişen sayıdaki eşlere verilir. Bu sayede anahtardan sorumlu eş ağdan çıkarsa yedek sorumlu eş bu sorumluluğu devralır. Yedek eş sayısı ağın ne kadar yüksek bir erişilebilirlik sağlayacağını belirleyecektir. DHT çözümü, çok büyük sistemlerin önbelleklemesi, karmaşık ve hızlı ölçeklenen sistemlerin koordinasyon çözümleri gibi günümüzün en yüksek sistem yüklerini koordine eden bulut mimarilerinde yaygın olarak kullanılan, kendini kanıtlamış bir güvenilir bir çözüm olarak karşımıza çıkmaktadır.

2.2. Makine Öğrenmesi Yöntemleri ve WEKA

Bu tez çalışmasında, makine öğrenmesi algoritmaların gerçekleştiriminde WEKA API [6] kullanılmıştır. WEKA, makine öğrenmesi çalışmalarında kullanılması amacıyla Waikato Üniversitesinde geliştirilmiş ve "*Waikato Environment for Knowledge Analysis*" kelimelerinin baş harflerinden oluşmuş yazılımın ismidir. Günümüzde yaygın kullanımı olan çoğu makine öğrenmesi yöntemini ve yöntemlerini, parametreleri değiştirilebilir, atanmış varsayılan değerleri ile içermektedir. Bunun yanında öznitelik seçimi sağlayan *Feature Selection, Attribute Evaluation* algoritmalarını desteklemektedir ve söz konusu yöntemler bu çalışma sırasında etkin bir şekilde kullanılmıştır. Ayrıca yöntemlerin sonuçlarını ölçme ve değerlendirebilmemizi sağlayan değerlendirme algoritmalarını da içermektedir. WEKA'nın hem doğrudan bir makine öğrenmesi masaüstü uygulaması, hem komut satırı üzerinden kullanımları mümkünken, biz WEKA'ya ait Java API'si ile benzetimimize bütünleştirmeye hazır şekilde kullanmayı tercih ettik. Aşağıdaki kesimlerde, tez çalışması kapsamında kullanılan makine öğrenmesi yöntemlerine kısaca değinilecektir.

2.2.1. C 4.5 Karar Ağacı

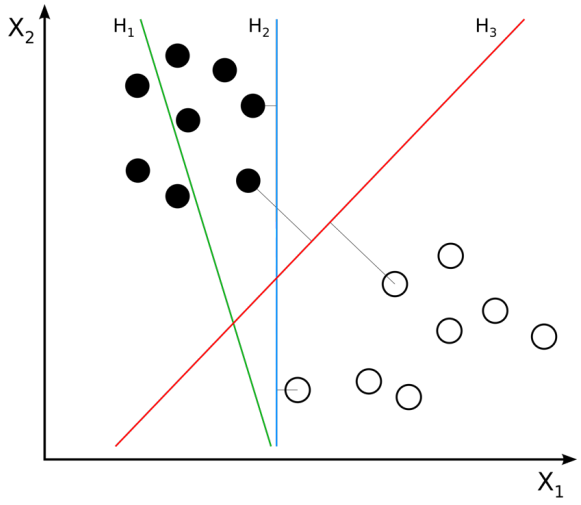
ID3 algoritması üzerine geliştirilmiş olan, karar ağacı inşa edilmesini sağlayan danışmanlı (*supervised*) bir algoritmadır. Bilgi kazanımı (*information gain*) ilkesine göre sınıflandırma yapan istatistiksel bir ağaçtır. Karar ağacını oluştururken, veri kümesini sınıflamak için en doğru öznitelik seçilerek ağaç büyütülmektedir.



Şekil 2.1 Örnek C 4.5 karar ağacı [7]

2.2.2. Support Vector Machine

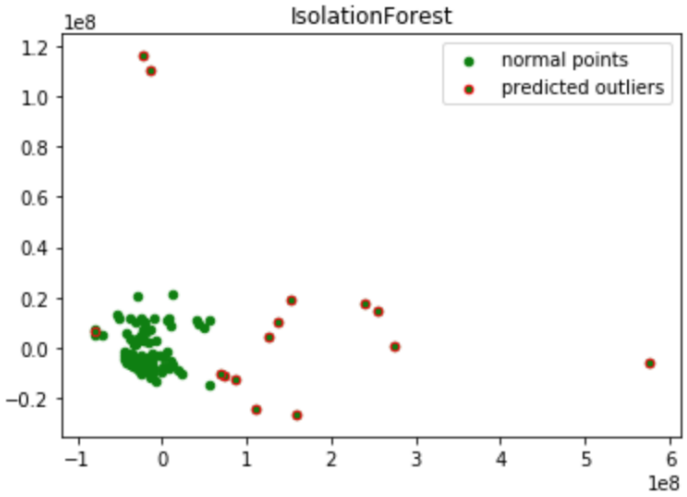
Regresyon analizi ile sınıflandırmayı sağlayan danışmanlı bir öğrenme algoritmasıdır. Değişkenler arası örüntülerin bilinmediği veri kümelerinde uygulanan istatistik temelli bir makine öğrenmesi yöntemidir. Algoritmaca belirlenen hipotezler içinde, sınıflar arası marjini en fazla artıran hipotez karar vermek hiperdüzlemi (hyperplane) olarak seçilir ve sınıflama bu hiperdüzleme göre yapılır. Örneğin Şekil 2.2'de H_3 hiperdüzlemi sınıflar arası marjini maximize ettiği için karar sınırı olarak kullanılmaktadır.



Şekil 2.2 Örnek SVM hiperdüzlemleri [8]

2.2.3. Isolation Forest

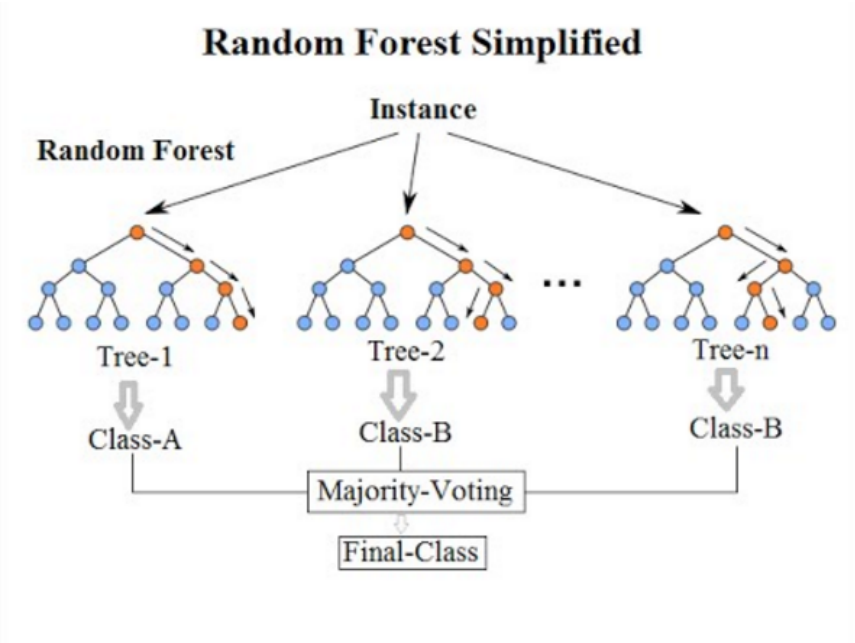
Pek çok yöntem normali tanımlayarak aykırı durumlar ile mücadele ederken, Isolation Forest aykırı durumu izole etme prensibi ile çalışan, danışmansız bir öğrenme algoritmasıdır. Isolation Forest bir noktanın normal olarak işaretlenmesine değil, herhangi bir noktanın diğer noktalardan ne kadar uzak kaldığına bağlı olarak aykırı nokta olarak işaretlenmesine karar verir. Bu sayede aykırı durumlar işaretlenerek izole edilmiş olur.



Şekil 2.3 Isolation Forest ile aykırı noktaların izolasyonu örneği [9]

2.2.4.Random Forest

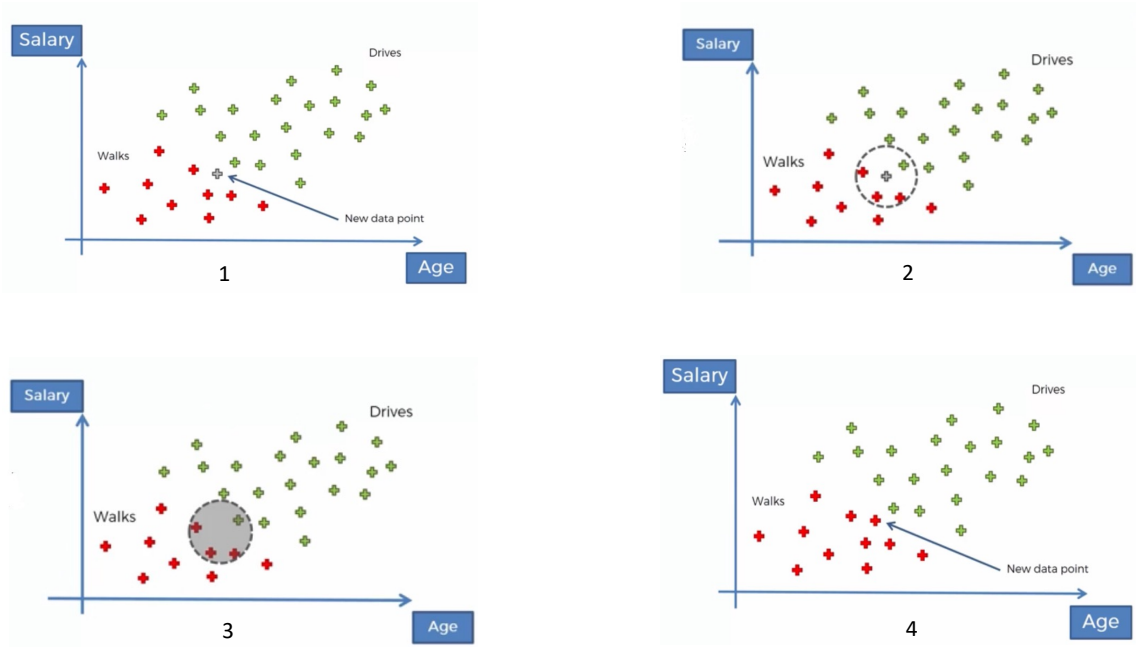
Random forest, sınıflandırma, regresyon ve diğer görevler için, öğrenme aşamasında çok sayıda karar ağacı oluşturarak problemin tipine göre sınıf veya sayı tahmini yapan bir toplu öğrenme yöntemidir. Yani tek bir karar ağacı yerine pek çok karar ağacı oluşturur. Ortalama bir sonuç elde ederek çıktı üreten Random Forest, değişken sayısının gözlem sayısından çok olduğu ortamlar için mükemmel sınıflandırmalar yaparken, diğer senaryolara kolayca uyarlanabilir olması ona geniş bir kullanım alanı sağlamaktadır [10].



Şekil 2.4 Örnek Random Forest çoklu karar ağacı [11]

2.2.5. Naive Bayes

Naïve Bayes sınıflandırıcı, örüntü tanıma problemine ilk bakışta oldukça kısıtlayıcı görülen bir önerme ile kullanılabilen olasılıksal bir yaklaşımdır. *Bayes* teoremi üzerine tanımlanmış, her bir özniteliğin istatistik olarak bağımsız değişken olması varsayımına dayanmaktadır.

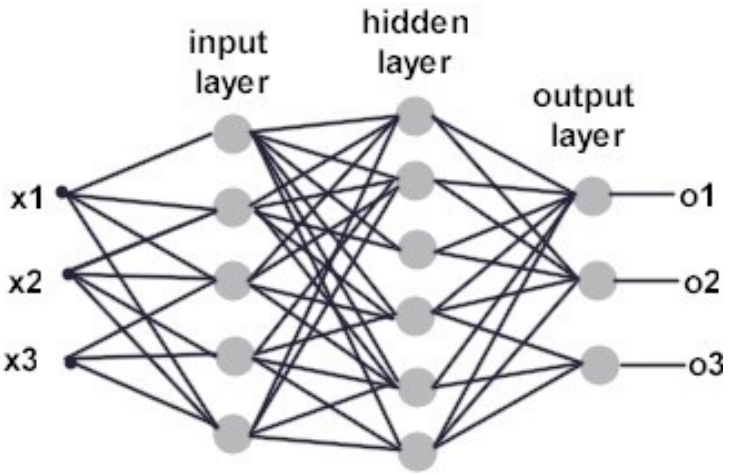


Şekil 2.5 Yeni kişi işe yürüyerek mi arabayla mı gidecek sınıflandırmasına ait Naive Bayes'in adım adım uygulaması [12]

Şekil 2.5'de dört adımda Naive Bayes'in bir sınıflandırmaya uygulanması gösterilmektedir. Birinci adımda sınıflandırılacak kişi konumlandırılıyor, ikinci adımda *Evidence Likelihood* hesaplanarak belli bir uzaklıktaki kişiler seçiliyor. Üçüncü aşamada seçilen kişilere *Likelihood* ile her bir noktanın hangi kümeye düşeceği olasılıkları hesaplanarak rasgele seçilen bir noktanın sınıfının ne olacağı belirleniyor. Son adımda ise *Posterior Probability* hesaplanarak yeni kişinin işe yürüyerek mi yoksa araçla mı gideceği olasılıkları arasından en yüksek olasılıklı sınıf seçiliyor. Şekil 2.5'de verilen senaryoda kişinin işe yürüyerek gidecek (Walk) olarak sınıflandırılmaktadır.

2.2.6. Multilayer Perceptron

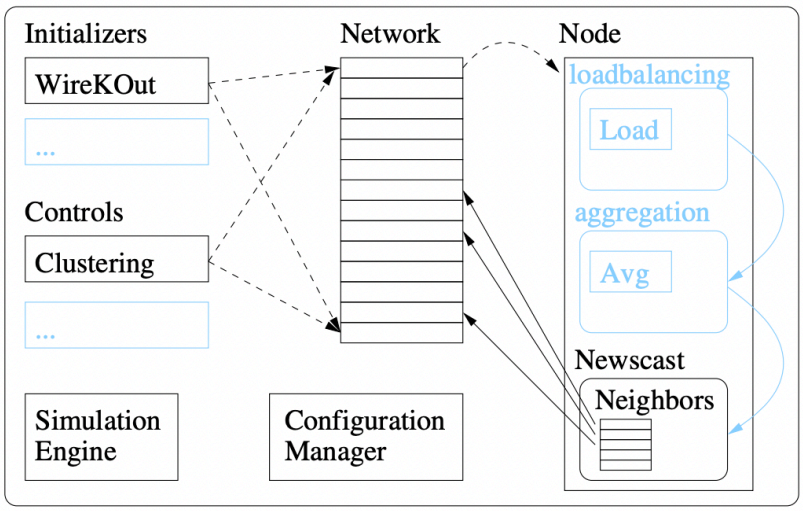
Yapay sinir ağları beyin ve sinir ağlarının yapısının taklit edildiği yaklaşımların tamamını çatısında toplayan paradigmadır. *Multilayer Perceptron* da örnek yapay sinir ağları modelidir. Sınıflandırma ve genelleme yöntemi olarak yaygın olarak kullanılan, yapay sinir ağlarında en çok bilinen yöntemlerden birisidir. Her katmanın çıktısı diğer katmanın girdisi olur. Son katmanda ise sonuca ulaşılır ve sınıflandırma elde edilir. Bir girdi katmanı, bir çıktı katmanı ve bu iki katman arasına yerleşmiş gizli katman ya da katmanlardan oluşur. Şekil 2.6'de MLP'nin basit bir gösterimi verilmiştir.



Şekil 2.6 Örnek bir MLP yapısı [13]

2.3. Peersim

P2P protokolünün kullanıldığı ağlara ait deneysel veri kümeleri yok denecek kadar az ya da yalnızca ağ trafiğinin benzetimlerinden oluşmaktadır. Bu alanda yapılan çalışmaların büyük bir kısmının sosyal ağlarla ilgili olması, söz konusu veri kümelerine erişimin kolay olmasından ileri gelmektedir. Oysa gerçek bir P2P protokolünde merkezi bir veri kaydı bulunmadığı için veriyi benzetimle elde etmek, bu alandaki çalışmalar için motivasyon kırıcı zahmetli bir iştir. *Peersim* [5] benzetim aracı söz konusu eksiğin giderilmesine yardımcı olmak için çıkarılmış, döngü tabanlı (cycle-based) ve olay tabanlı (event-based) benzetim yeteneğine sahip, açık kaynak bir araçtır.



Şekil 2.7 Peersim üzerinde örnek bir gerçekleştirime ait bileşenler [5]

Peersim Java ile geliştirilmiştir ve yapılandırılabilir olmasının yanında, kaynak kodlar üzerinde kolayca genişletilerek yeni yetenekler kazandırmaya uygundur. Şekil 2.7'de verilen örnek bileşenler, Peersim üzerinde Newcast [14] gerçekleştirimine ait bileşenlerdir. Peersim ile kişisel protokoller, akışlar, denetimler ve başlatıcılar (*initializer*) geliştirilebileceği gibi, kullanıma hazır pek çok bileşenden faydalanılabilir olduğunu söyleyebiliriz.

3. İLGİLİ ÇALIŞMALAR

Geleneksel olarak güven yönetimi yaklaşımları eşler arası ağın yapısına göre iki sınıfta değerlendirilir. DHT temelli yaklaşımlarda [15-21], eşe ait genel güven bilgisi DHT mekanizması tarafından belirlenen bir başka eş tarafından saklanır. Böylece eşe ait tarihsel etkileşimler, DHT üzerinde eşle ilgili kayıtları tutan başka bir eş üzerinden erişilebilir olur. Yapısal olmayan yaklaşımlara sahip pek çok çalışmada ise sistemler güven yönetimi ilgili eşe ait yerel güven bilgisinin eşler arasında yayılımına dayanmaktadır. Yapısal olmayan ağlarda eşlere ait bilgiler komşu eşlerde ya da daha önce etkileşime geçilmiş eşlerde tutulur [22-27]. Güven değerini sorgulayan bir eş, sorgusunu komşularına aktarır, komşuları da kendi komşularına aktararak bu şekilde bilgi talebini yayar. Bu sıçrama sayısı sorgu alanına ya da ağ gereksinimlerine göre yapılandırılabilir. Bu tür bir akışa dayanan güven hesaplaması çoğu zaman tüm görüşlerden yararlanamadan, sadece erişilen eşlerin temin ettiği verilere dayanarak yapılır.

3.1. Geleneksel Yöntemler

Güven yönetimi alanında çalışmalar uzunca bir süre ağırlıklı olarak istatistiksel modellere dayanıyordu. Özellikle DHT tabanlı çoğu yöntem Aberer'in modeli ile [15] Eigentrust [16] yönteminden esinleniyordu. Aberer'in modeli güven yönetimi ve itibara dayalı güven yönetimi konseptine giriş yapan ilk çalışmalardandı. Bu model, yeni katılan bir iş ile eski bir eş arasında bir fark gözetmeksizin, P-Grid yapısı üzerinde olumsuz deneyimleri paylaşmaya dayanıyordu. Bu da onu aklama (*whitewashing*) saldırılarına karşı zayıf kılıyordu. Eigentrust modeli [16] yerel ve global güven değerleri hesaplarken, bu hesaplamayı dağıtık ve merkezi olmayan bir model üzerinde hesaplanan Eigen vektör hesaplamasına dayandırır. Eigentrust modeli hem olumlu hem olumsuz deneyimlerin paylaşımına dayandığından daha kararlı bir sonuç verir. Xiong ve Li, PeerTrust [17] ismini verdiği modelinde geri bildirim güvenilirliği (*feedback credibility*) konseptini uygulayarak bu alanda benzerlik kullanımına güçlü bir inanç oluşturan bir çalışma oldu. Bayes yöntemi ile eşleri tanımlayan ve sınıflandıran bir yöntem uygulayan Wang ve Vassileva [28], hizmetleri film, müzik gibi kategorilere ayırırken hizmet kalitesini de indirme hızına göre yavaş, normal ve hızlı olarak 3 kategoriye ayırdı. Her bir ana kategori üzerinden memnuniyetin (satisfaction) puanlandığı model eşlere belli bir kategoride belli bir kalite seviyesinde hizmet almalarını sorgulayabilmelerini sağladı. Oldukça geleneksel

bir yöntem olan çalışma güven yönetimine bağlam ekleyen ilk çalışmalardan olarak karşımıza çıkıyor. Swamynathan ve diğerleri çalışmalarında [29] eşlerin hizmet sağlayıcı ile hizmet değerlendirici olarak ayrıldığı modelde *geri bildirim ağırlığı* verilen geri bildirim ile çarpılarak normalize ettiler. Vektör paylaşımına dayanan Gossiptrust [25] modelinde Zhou ve diğerleri vektördeki değerleri ikiye bölüp, bir yarısını diğer eşlerden gelen görüşlerle yeniden birleştirerek güven değerine ulaşmaya çalıştılar. Bu çalışmada global skor saklama ve getirme için *Bloom Filter* kullanılmıştır. Chen ve diğerleri topoloji temelli bir güven yönetimi sundu [30] ve modelini iki temel bileşenden oluşturdu; itibar mekanizması (*Reputation mechanism*) ile küme itibar mekanizması (*Cluster Rep Mech*.) ağdaki trafik yükünün önemsendiği modelde kümeler arası iletişimi sağlayan geçit, küme içi eşleri yöneten küme lideri (*cluster head*) gibi konseptler bulunuyor. Hesaplamalarda zamana önemli bir bileşen olarak yer verilen modelde, sistemde geçirilen sürenin artırılması hedeflendi. İtibar hesaplamalarına göre bencil davranışlar cezalandırılırken, iyi hizmete teşvik edilmesi de amaçlanmıştı. İnsanların kurduğu sosyal ilişkilerden öykünerek kurulan bir topoloji ile güven yönetimine başvuran Han [31] *Semantic Overlay'*i insan ilişkilerinde olduğu gibi taklit etti. Modelde bir topluluğa (*community*) dahil olma süreci onaya bağlı iken, girilen bir toplulukta yer alan her bir eşin davranışları tüm topluluğu etkiledi. Çalışmada üç temel güven bileşeni önerildi; davranış kimliği (*behavior credential*), topluluk kimliği (*community credential*) ve itibar kimliği (*reputation Credential*). Bunların toplamından güven değerine ulaşılan modelde topluluk içi sadece davranış kimliğine bakılırken, topluluklar arası güven hesaplaması için ise tüm bileşenlere başvuruluyor. Güven ve itibar kavramlarına başvuran Liu [20] çalışmasına da konu olan iki temel kavramını şöyle tanımlıyor: itibar değeri eşin kaliteli hizmet sağlayıp sağlamadığını temsil ederken, güven değeri eşin kendisine hizmet sağlayan eşe adil puan verip vermediğini temsil eder. SORT [22] modeli tarihsel veriler ve komşu geri bildirimleri ile güven ilişkilerini yönetti. Bu çalışmada hizmet ve itibar bağlamları tanımlandı ve bir hizmet tatmin, ağırlık ve zamana dayalı bozunma etkisi üzerinden değerlendirildi. Cornelli ve diğerleri [9], talep sahibinin ve sağlayıcının aykırı durum olma durumunu korurken dağıtık bir seçim algoritmasına dayalı itibar paylaşımı modeli önerdi. Geleneksel yöntemlerin sosyal ilişkilerden esinlendiği pek çok çalışma da bulunmaktadır. Sorcery çalışmasında Zhai ve diğerleri. [32] sosyal ilişkilerden elde edilecek *baskın bilgi* sayesinde bir eşin bir başka eş ile ilgili yalan söyleyip söylemediğinin tespit edebileceği bir model önerdi. Bir başka geleneksel sosyal ağ temelli güven yönetimi modeli olan SocialLink [33], potansiyel kötücül davranış etkileşimlerini

önlemeye çalışırken, eşlerin, arkadaşı olmayan eşlere de hizmet vermesine teşvik eder ve yalnızca hizmet almaya çalışan (free rider) eşleri engellemeyi hedefler.

3.2. Benzerlik Temelli Yöntemler

Benzerlik kavramı farklı isimlendirmelerle hem geleneksel yöntemlerde hem modern yöntemlerde karşımıza çıkabilmektedir. Geleneksel yöntemler içerisinde PeerTrust dışında da pek çok çalışmaya konu edilmiştir. FCTrust [21] eşlerin güvenirliğini ölçmek ve değerlendirmek için geri bildirim güvenilirliğine dayalı, dağıtık bir eşler arası global güven modelini kullanır. Eşlerin birbirlerine olan güvenini ise eşler arasındaki etkileşim sıklığı ile diğer eşlere verilen geri bildirimlerin birbirine benzerliği üzerinden hesaplar. ServiceTrust'ta [26] benzerlik pozitif benzerlik (*similarity*) ve negatif benzerlik kavramı olarak ikiye ayrılarak olumlu görüşler ile olumsuz görüşlerin etkilerinin daha parametrik yönetilir olması sağlandı. Devam çalışması olan ServiceTrust++ modelinde [27] Su ve diğerleri, modele bozunma faktörü (*decay factor)*, benzerlik, eşik, kontrollü rastgelelik ve atlama stratejisini dahil etti. Xei ve diğerleri [34] önerdiği STTM modelinde ortak komşular ile etkileşimlerin benzerliği kosinüs benzerliği ile hesapladı. STTM modelinde yerel güven değeri için bu benzerlik kullanılırken global güven değeri için MLE (*Maximum Likelihood Estimation*) kullanıldı. Guo çalışmasında [19] zamana duyarlılığı ekledi ve davranışların da hassasiyetle tespitini hedefledi. Benzerlik için yeterince veri olmadığında MLE yönteminin kullanıldığı modelde, iki eş arasındaki etkileşime tanıklık etmesi için tanık eş *(*witness peer) kavramına yer verdi. Das ve Islam geçmiş güven (*Past Trust*), güncel güven (*Recent Trust*) *ve* bozulma modeli (*Decay Model*) gibi zamanı önemli bir boyut olarak çalışmasına dahil etti [35]. Güven yönetiminde eşlerin kendi yerel güven değerine ve genel itibarlarına, benzerlik ve güvenilirlik ile güçlendiren model tüm bu parametreler ile nihai bir güven değerine ulaşıyor. Modelin temel motivasyonu davranışını zekice değiştiren saldırganlara karşı hızlı adapte olan bir model oluşturmaktı. İki eş arasındaki benzerliğin dikkate alındığı bir başka çalışmasında Prasad [36], kendisine benzeyen eşlerin birbirlerine daha yüksek itibar verdiği bir model oluşturdu. Bu modelde etkileşim sayısı, sıklığı, aynı eşlerle etkileşim ve farklı eşlerle etkileşim sıklıkları, etkileşimlerin ne kadar süre önce gerçekleştiği gibi parametreleri dikkate alarak bu benzerlik hesaplandı.

3.3. Gelişmiş İstatistiksel Yaklaşımlar

Geleneksel istatistiksel yöntemlerin yanı sıra HMM (*Hidden Markov Model*) uygulamaları ve bulanık mantık uygulamaları güven yönetimi için sıkça başvurulan yöntemler olarak karşımıza çıkmaktadır. Güven hesaplama sürecini göreli belirsizlik ölçümü olarak değerlendiren Ouyang ve diğerleri [37] geçmiş verilere dayanan bir olasılıksal belirsizlik çözümlemesi için olasılık teorisi ve istatistik kullanmış, HMM sürecini hızlandırmak için *Bonus Malus System*'den faydalanmıştır. Müşteriye göre otomobil sigortası fiyatlandırma yöntemi olarak kullanılan *Bonus Malus System'*i modellerine uyarlayarak, HMM parametrelerinin tahminleme maliyetini düşürerek modeli basitleştirmeyi hedeflemişlerdir. Yahyaoui [38] web servislerine odaklanan çalışmasında gelişmiş bir HMM örüntü temelli güven önyükleme modeli önerdi. Bu çalışmada önce bir web servisinin tanımlanmasını sağlayacak güven örüntülerinin özellikleri çıkarılıyor; geliştirilen model ile web servislerinin davranışlarının bu ön tanımlı güven örüntülerinden birisi ile eşleştirilmesi sağlanmıştır. Video içeriği sağlayıcı hizmetler için adanmış bir eşler arası ağda hizmet kalitesini artırmayı hedefleyen modelinde Huang ve diğerleri [39], davranışları kaytarma (free-riding) yoğunluğuna göre sınıflandırarak kötücül davranış olarak yoğun frekansta kaytaran eşlerin davranışı belirledi. HMM ile itibarın hesaplandığı modelde kaytaran ile kötü niyetli eşleri ayırt etme algoritması önerilmiştir. Tian ve Yang çalışmasında DS (*Dempster-Shafer*) kanıt teorisi üzerine kurulan bulanık mantık kurallarını eşler arası ağlarda güven yönetimi için kullanarak, özellikle yönlendirici bilgilerin çelişkili ve tutarsız olduğu senaryolarda başarılı sonuçlar elde ettiğini not etti [40]. Bir diğer çalışmada bulanık mantık ile kurulan itibar modellerinde eş sayısı arttığında hesaplama maliyetinin üstel olarak arttığını tespit eden Guo ve diğerleri, bulanık kümelemeye dayalı maksimum ağaç yöntemi ile eşleri sınıflandırarak hem başarılı bir sınıflandırma elde etmiş hem de hesaplama maliyetlerini ciddi anlamda düşürmüştür [41]. Saeed ve diğerleri çalışmalarında eşler, kaynakları arasındaki anlamsal benzerliğe göre kümelere ayırdı [42], bu benzerliğe göre eşler arası ağdaki konumu belirledi. Ayrıca çalışmada açık sözlülük gereksinimi ve genişleyebilir olma yeteneğini sağlamak için OWL (*Web Ontology Language*) dilinden faydalanılmıştır. Dempster-Shafer kullanılan bir başka çalışmada Feng güven değerlendirmelerinin geçişliliği ile dinamik birleştirme ile güven önerisini hesaplamaktadır [43].

3.4. Makine Öğrenmesi Yaklaşımları

Temelde geçmiş etkileşimlere dayanan geleneksel yöntemlerin "soğuk başlangıç" ve "sıfır bilgi" sorunlarının üstesinden gelmek için makine öğrenimi yöntemleri güven yönetiminde son yılların daha çok tercih edilen yaklaşımı olarak karşımıza çıkmaktadır. Yapay zekanın da önemli bir alt başlığı olan makine öğrenmesi, yeterince veriye dayandığında, öznel bir değerlendirme de içeren güven sınıflandırması için geleneksel yöntemlere göre daha insansı sonuçlara ulaşılabileceği düşünülebilir. Makine öğrenmesi güven hesaplamalarında olağan hali ile saldırı ya da saldırgan sınıflayıcı olarak konumlandırılmaktadır. İstisnai olarak saldırıyı tespit eden değil tespiti destekleyen çalışmalar karşımıza çıkabilmektedir [44].

Yaygın bir ortamda güven yönetimini dinamik olarak ele alacak bir model öneren Yuan ve diğerleri önceki olasılık, güven değeri, geçmiş etkileşimler, zaman etkisi ve eş itibarlarını etkenler olarak tanımlayarak NBC (Naive Bayes Classifier) algoritmasını uyguladı [45]. Algoritmanın iki kez uygulandığı modelde birinci karar yalnızca hizmet sağlayıcının önceki bilgisine göre karar verirken, sonuç elde edilemez ise algoritmanın ikinci uygulanışında son kararın verilmesi için öneri bilgisi parametre olarak eklenerek hesaplamalara dahil edilir. NBC algoritmasından yararlanan bir başka çalışmada D'Angelo [46] *priori* algoritması ile NBC temelli bir güven modeli önerdi. Priori algoritması ile dokuz özellik (*property*) çıkaran model NBC algoritması ile veriyi güvenilir, şüpheli ve güvenilmez olarak sınıflandırıyor. Baohua ve diğerleri BP-NN (*Back Propagation Neural Network*) yöntemini kullanarak eşe ait yerel güven değerini hesaplayan bir model önerdi [47]. Büyük dağıtık sistemlerde güven mimarisi öneren Liu ve diğerleri çalışmasını bilgi toplayıcı (knowledge collector), güven hesaplama motoru ve saklama bileşeni olmak üzere üç bileşen ile modelledi [48] ve makine öğrenimini geçmiş etkileşimleri başarılı/başarısız sınıflandırması için kullandı. Kesikli geri bildirimleri Q-Learning algoritması ile dinamik olarak ele alan Aref ve Tran çalışmasında bulanık mantığı güven tanımının belirsizliğini ve kesin olmamasını ele almak için makine öğrenmesi ile birleştirdi [49]. HMM'yi kullanarak güven örüntüleri çıkaran Yahyaoui [38] ilk çalışmasında 5 kategori tanımlarken, devam çalışmasında [50] bu sayıyı 11'e çıkararak kural tabanlı Prefix-Suffix algoritmasını uyguladı. Benzerlik kavramını makine öğrenmesi yöntemlerinde kullanan Korovaiko ve Thomo çevrimiçi toplulukların kullanıcıları arasındaki güveni modelleyen bir yöntem önerdi [51]. Modelinde kullanıcı

benzerliğini işaret eden beş ile değerlendirici görüşlerini işaret eden üç öznitelikten faydalandı. Mao, sinir ağlarını kullanarak hizmet kalitesi ile hizmet güvenilirliği arasındaki doğrusal olmayan ilişkiyi çıkararak bulut hizmetlerin güven değerini değerlendirdi [52]. IOT ortamında güven yönetimi için makine öğrenmesinden yararlanan Jayasinghe ve diğerleri çalışmasında [53] ilişkiler, konum, itibar gibi bilgileri kullanarak danışmansız (unsupervised) öğrenme ile etiketleme yaptı. Elde ettiği veriyi danışmanlı (*supervised*) öğrenme için kullanarak değerlendirme yapan bir model önerdi. Hizmet kalitesine ait nitelikleri kullanarak web servislerinin güven değerlendirmesini yapan bir model öneren Ramakanta ve diğerleri çalışmasında değerlendirme sürecini üç adımda gerçekleştiren bir model önerdi [54]: öznitelik seçimi(1), sınıflandırma modeli öğrenimi(2) ve kural üretimi(3).

Sosyal ağlarda güven yönetimi alanında makine öğrenimi kullanımı diğer alt alanlara göre ön plana çıkmaktadır. Eşlere ve eşler arası etkileşimlere ait öznitelikler ile ML uygulayan Liu ve diğerleri güven faktörünü de eşlere ve etkileşimlere ait faktörler olarak ayrı değerlendirdi [55]. Etkileşimler üzerinden elde edilen sonuçların eşler üzerinden elde edilen sonuçlardan daha iyi sonuçlar verdiğini not etti. Zolfaghar ve Aghaie beş adımda açıkladığı tahminleme sürecinde güven uyandıran (trust-inducing) faktörleri ile güvene dayalı ağların etkileşim verilerindeki seyrekliği (sparseness) adresleyen bir çatı önerdi [56]. Çalışmasında Facebook kullanıcıları arasındaki güven ilişkileri üzerinde çalışmalar yapan Yuji öznitelikler ile güven değeri arasındaki Pearson korelasyonu sabiti üzerinden etkileşimler ve profiller bilgilerinden öznitelikleri seçerek KNN, SVM ve MLP yöntemlerini uyguladı [57]. Papaoikonomou ve diğerleri [58] çalışmasında iki kullanıcı arasındaki güven seviyesini yine kullanıcıların değerlendirmelerini yarı-danışmanlı (*semi-supervised*) öğrenme yöntemi elde eden bir model önerdi. Bu çalışmada *Autoencoder* ile tüm veriye danışmansız öğrenme uygulandıktan sonra sinir ağları ile sınıflandırdı. Chen, kullanıcı özniteliklerine uygulanan makine öğrenmesine dayalı bir güven değerlendirme çatısı (framework) önerdi [59]. Twitter verilerine uygulanan çalışmada öznitelikler dört kategoriye ayrıldı ve öznitelik seçim algoritmaları uygulanarak güven kararına ulaştı. Yayılma (*propagation*) ve pekiştirmeli (*reinforcement)* öğrenme ile güven değerlendirme yöntemi öneren Kim ve Song, kısa mesafe doğrudan ilişkiler üzerinden uzun mesafe doğrudan olmayan ilişkileri tahminlemeye çalıştı [60]. Q-learning aşamasını Markov karar oluşturma sürecine

dayandırdı. Kümeleme algoritması ile Mobil Sosyal Ağlar (MSN) için geri bildirimlerde demet (*tuple*) kullanan bir güven değerlendirme modeli öneren Chen [44] doğrudan güven kararı vermek yerine kararı destekleyecek sonuçlara odaklandı.

3.5. Farklı Alanlarda Güncel Güven Yönetimi Çalışmaları

Son yıllarda güven yönetiminin sıcak başlıkları olarak blok zinciri ve IOT üzerinde yapılan pek çok çalışma yayınlandı. Barenji blok zinciri temelli güven yönetimi modeli "Blocktrust" ile bulut hizmetlerindeki güven sorununa çözüm aradı [61]. DTMS (*decentralized trust management system)* modeli [62], merkezi olmayan bir fikir birliğine dayalı güven değerlendirmesini, şeffaf bir değerlendirme prosedürü ile ele almayı hedefledi. Güven kredilerinin geri döndürülemez şekilde depolanmasını sağlayan blok zinciri tabanlı bir güven depolama sistemini benimsedi. Abdelghani [63] ve diğerleri, özellikle Sosyal IoT ortamları için tasarlanmış karmaşık, kısıtlı ve son derece dinamik ağların üstesinden gelmek için yeni bir dinamik ve ölçeklenebilir çok seviyeli güven modeli olarak DSL-STM modelini önerdi. Wu ve Liang [64] blok zinciri ve IOT'yi bir araya getirdi ve algılayıcı (sensor) düğümlerinin güvenilirliğinin mobil uç düğümler tarafından değerlendirildiği blok zinciri tabanlı bir güven yönetimi mekanizması (BBTM) önerdi.

4. TUTARLILIK TEMELLİ İSTATİSTİKSEL MODEL

Tutarlılığa (*consistency*) dayalı temel modelimizde her bir eş hizmet ya da kaynakları sisteme sunarken, diğer eşler hakkındaki güven hesaplamasını etkileyecek bilgileri de saklar. Güven bilgileri, tüm eşlerin etkin bir şekilde ulaşabilmesi için DHT tabanlı bir ağ üzerinde tutulmaktadır. Her bir eşin ve kaynağın biricik kimliğe (ID) sahip olduğu varsayılmıştır. Her bir eşin geçmişi, kendisinden sorumlu Arşivci (*Archiver*) rolü ile tanımlanan eşlerde saklandığı için herhangi bir eşe ait geçmiş bilgiler değiştirilemez veya silinemez. Bir etkileşimdeki hizmet sağlayıcı için hizmet alan eşin yaptığı değerlendirmeye geri bildirim denir. Tez kapsamında, bir geri bildirimin belli bir bağlamda diğer geri bildirimlere olan benzerliğini tutarlılık olarak tanımlıyoruz.

4.1. Arşivci

Pek çok modelde hizmet alan ve hizmet veren eşler dışında farklı rolleri üstlenen roller zaman zaman başvurulan bir yöntemdir. Bir küme içinde küme lideri olabildiği gibi [30] etkileşime şahitlik etmesi için tanımlanan roller de olabilmektedir. Bu çalışmada, eşler için hizmet alan eş veya hizmet veren eş olmak dışında bir rol tanımlanmıştır.

Bir eşe ait tüm bilgiler, o eşe ait özel bir arşivcide tutulur. Bir eşe (bu eşe x diyelim) ait değerli bilgiler:

- $\mathrm{F}_p(x)$: Bir eşin hizmet sağlayıcı olduğu etkileşimlerde aldığı *geri bildirimler*
- $\mathrm{F}_r(x)$: Bir eşin hizmet alıcısı olduğu etkileşimlerde verdiği *geri bildirimler*
- Eşe ait tutarlılık $(PC(x))$ ve güven değerleri $(T(x))$
- Tamamlanmış ve devam eden tüm etkileşimler

Her bir geri bildirim, paketlenmiş bilgi demeti (*tuple*) olarak tutulur. $f_i(x,y) = (s_i(x,y), FC_i(x,y))$ gösterimi hizmet sağlayan bir x eşine, hizmet alan y eşi tarafından verilen i sıra numaralı etkileşimin (*interaction*) geri bildirimi olarak ele alınmakta, $s_i(x,y)$ gösterimi memnuniyet (*satisfaction*) değerini ve $FC_i(x,y)$ geri bildirim tutarlılığı bilgisini temsil etmektedir. Bir etkileşim başarılı bir şekilde tamamlanabilir, bir şekilde yarıda kalabilir (bağlantının kesilerek eşin çevrim dışı olması gibi) ya da hizmet sağlayıcı saldırı gerçekleştirebilir. Bu senaryoların her biri için hizmet alan ve etkileşim sonunda geri bildirim sağlayan eş şu şekilde davranır:

$$s_i(x,y) = \begin{cases} 1, \text{etkileşim başarılı ise} \\ 0, \text{etkileşim tamamlanmadıysa} \\ -1, \text{x kötücül bir davranış sergilediyse} \end{cases} \quad (1)$$

Arşivci rolüne sahip bir eş, bu görevini yerine getirirken yanlış bilgilendirmede bulunabilir. Modelimizin etkileşim döngüsünde buna önlem olarak, her bir eş için üç arşivci tanımladık ve her bir sorgunun çapraz doğrulamasını yaparak buna çözüm getirdiğimizi varsaydık. Bu sebeple deneysel çalışmalarda arşivci saldırıları ele alınmamıştır.

4.2. Geri Bildirim Tutarlılığı

Hizmet alan eş, hizmet sağlayan eş ile etkileşimi sonlandığında, girdiği etkileşimle ilgili değerlendirmesini geri bildirim olarak hizmet sağlayan eşin arşivci eşine gönderir. Bu geri bildirim tutarlılığı arşivci tarafından hesaplanarak geri bildirim tutarlılığı olarak kaydedilir. Geri bildirim tutarlılığı, bir eş ile ilgili verilen bir geri bildirimin, geçmiş geri bildirimler ile ne kadar benzediğini ölçer. Benzerlik ölçümü literatürde daha çok vektör-temelli karşılaştırma [19,26,34,65] ile hesaplanırken, bu yöntem bizim yaklaşımımıza uymamaktadır. Bizim yaklaşımımızda yalnızca bir değerin vektör içerisindeki değerler ile benzerliğinin ölçülmesi gerekmektedir. Bir başka deyişle, geri bildirim tutarlılığı, değerlendirilen eş ile ilgili verilen bir geri bildirimin, o eşe verilmiş geçmiş geri bildirimler ile benzerliğidir. Aynı değere sahip geri bildirim sayısı geri bildirim tutarlılığını belirler. x'in y'ye hizmet sağladığını varsayalım. x eşine ait i numaralı etkileşim y eşi ile yapılıyorsa, x eşinin arşivci eşi (A_x) geri bildirim tutarlılığını aşağıdaki formül ile hesaplar:

$$FC_i(x,y) = \frac{[\mathrm{F}_p(x) \cap s_i(x,y)]}{[\mathrm{F}_p(x)]+1} \quad (2)$$

$[\mathrm{F}_p(x)]$ ifadesi $\mathrm{F}_p(x)$ içerisindeki geri bildirim sayısını, $[\mathrm{F}_p(x) \cap s_i(x,y)]$ ifadesi aynı memnuniyet değerine sahip geri bildirim sayısını, $s_i(x,y)$ i numaralı etkileşimdeki memnuniyetini ifade etmektedir.

4.3. Eş Tutarlılığı

Eş tutarlılığı doğru geri bildirim yapıldığının ölçüldüğü bir tanımlamadır. Bir hizmet alıcısı olarak y eşinin geçmiş geri bildirimleri dikkate alındığında, $\mathrm{F}_r(y)$ içerisindeki tüm

geçmiş geri bildirimlerin geri bildirim tutarlılığı değerleri, eş tutarlılığının bir ölçümü olarak kabul edilebilir. Böylece *y* eşinin eş tutarlılığı şu şekilde hesaplanır:

$$PC(y) = \frac{\sum_{f_i(*,y) \in \mathbb{F}r(y)} FC_i(*,y)}{[\mathbb{F}r(y)]} \quad (3)$$

Burada $f_i(*, y)$, y eşinin verdiği i numaralı geri bildirimi ve $FC_i(*, y)$ ise bu geri bildirime karşılık gelen geri bildirim tutarlılığı değeridir.

Geri bildirim tutarlılığı bir eş ile ilgili verilen önceki geri bildirimlere benzerliği ölçerken, eş tutarlılığı bir eşin hangi seviyede tutarlı geri bildirimler verdiğini ölçmektedir.

4.4. Güven değeri hesaplama

Bir eşin güven değeri, o eşe hizmet sağlayıcı olduğu etkileşimler sonucu verilen geri bildirimleri değerlendiren arşivci tarafından hesaplanır. Bir geri bildirimin değerlendirilmesi sırasında, o geri bildirimin geri bildirim tutarlılığı ve geri bildirimi veren eşin *eş tutarlılığı* dikkate alınır. Arşivci eş, kayıtlarını tuttuğu eş ile ilgili gelen yeni bir geri bildirim sonrası güven değerini hesaplayarak tutar. Bir *x* eşinin *i* numaralı hizmetini *y* eşine verdiğini, *y* eşinin etkileşime ait $f_i(x, y)$ geri bildirimini A_x 'e gönderdiğini varsayarsak; *x* eşinin güven değeri A_x tarafından şu şekilde hesaplanır:

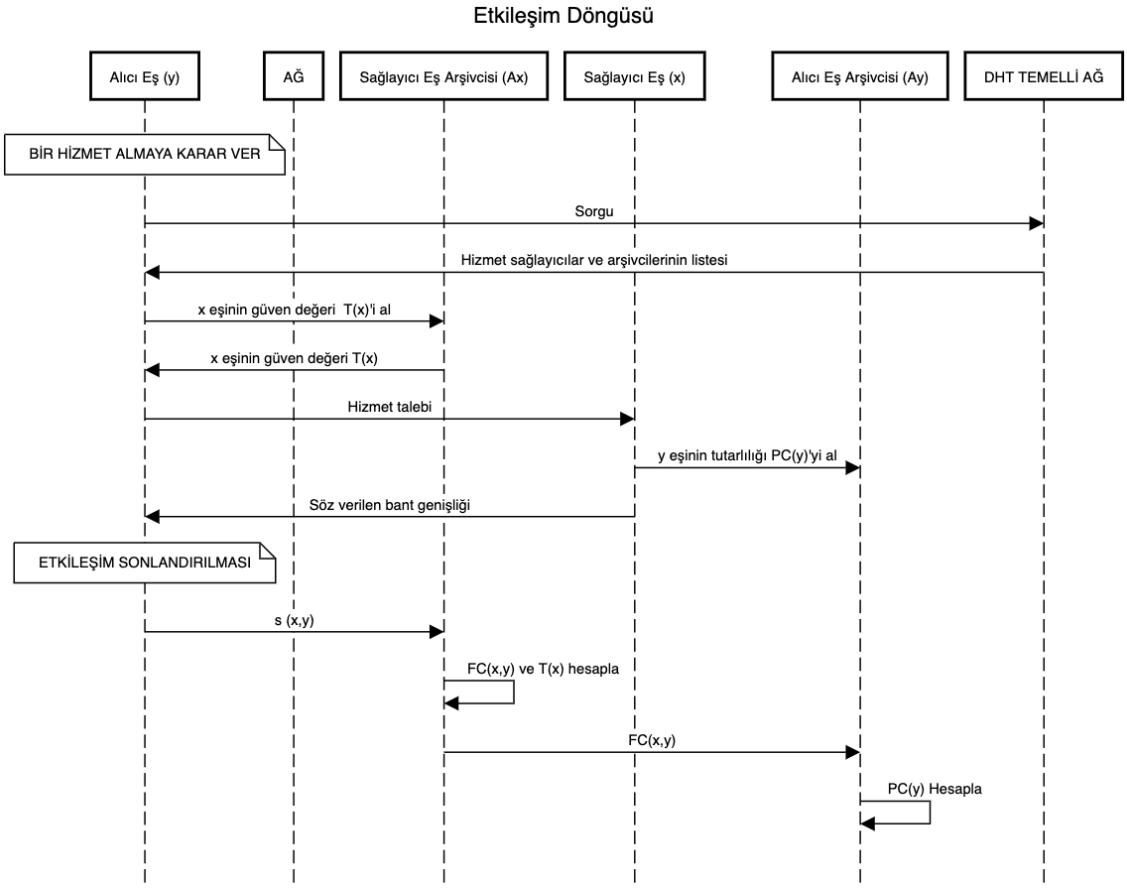
$$T_i(x) = \propto . E_i(x, y) + (1 - \propto) . T_{i-1}(x) \quad (4)$$

$$E_i(x, y) = s_i(x, y) . FC_i(x, y) . PC(y) \quad (5)$$

$T_i(x)$, *x* eşinin *i* numaralı etkileşimi sonrası güven değerini, $E_i(x, y)$ *x* eşine ait *i* numaralı geri bildirimin değerlendirmesini ve $0 < \propto < 1$ ise söz konusu geri bildirimin güven değerine etkisini belirleyen sabit değerdir. $E_i(x, y)$ hesaplanırken geri bildirim tutarlılığı ($FC_i(x, y)$) ve geri bildirim yapan eş tutarlılığı (örneğin hizmet alan *y* eşinin tutarlılığı $PC(y)$) dikkate alınır. Bu sayede bir geri bildirim ve onu sağlayan eş ne kadar tutarlı ise söz konusu geri bildirimin etkisi o kadar etkili olacaktır. Ağın oluşturulması, ilk değerlerin oluşması ve eşlere geçmiş verisi olmadan etkileşim şansı vermek için, her bir *x* eşine ilk güven değeri olarak $T_0(x) = 0.2$ atanır. Ayrıca, geçmiş geri bildirimlerin etkisi ile yeni bir geri bildirimin etkisini dengeleyebilmek için $\propto = 0.2$ olarak belirlenmiştir. Bu değerler kapsamlı deneyler yapılarak seçilmiştir.

4.5. Etkileşim döngüsü

Şekil 4.1, modelimizde etkileşimin nasıl gerçekleştiğini göstermektedir. Bir etkileşimin başlangıcındaki ilk adımında (örneğin dosya indirme), hizmet almak isteyen *y* eşi ağda olası kaynak sağlayıcıları sorgular. Sorgu sonucu ağdan hizmet sağlayıcı eşler ile onların arşivci eşleri elde edilir. Bu çalışma kapsamında sorgulanan kaynağın tüm sağlayıcı eşlerinin listesine tek sorguda ulaşılabildiği varsayılmıştır. Bunun yanı sıra, bazı ağ altyapıları tüm potansiyel sağlayıcı adayları yerine yalnızca bir alt kümesini bu sorgu sonucu olarak dönebilir. Ancak bu bizim hesaplamalarımızı ve matematiksel gösterimlerimizi etkilemeyecektir. Anlaşılır olması için sorgu sonucu dönen eşler arasında yalnızca bir *x* eşi ile onun arşivcileri arasından A_x gösterime eklenmiştir.



Şekil 4.1 Etkileşim döngüsü

İkinci adımda A_x, *x* eşine ait güncel güven ($T(x)$) değerini *y* eşine bildirir. Eğer $T(x)$ değeri belirlenen bir eşik değerinden yüksek ise *y* eşi hizmet sağlayıcı adayı olarak *x* eşine hangi büyüklükte bir bant genişliği sağlayabileceğini üçüncü adımda sorar. Hizmet sağlayıcı elemek için kullanılan eşik değeri (*threshold)* ilk olarak 0,8 olarak atanır. Eğer

bu eşik değerini aşan bir güven değerine sahip eş bulunamazsa ya da bulunan eşler *y* eşini reddederse eşik değeri kabul eden bir hizmet sağlayıcı eş bulunana kadar sırası ile 0,6, 0,4 ve 0,2'ye kadar indirilir. Eşik değeri 0'a kadar indiğinde hizmet sağlayıcı eş arama işlemi durdurulur ve istek iptal edilir. Eşik değeri kademeli olarak düşürülerek hizmet sağlayıcı bulma şansı artırılmış olur. Fakat, sistemde daha güvenilir bir etkileşim ortamı sağlanmak istenirse eşik değerinin en düşük hedefi 0.2 yerine daha yüksek bir değer seçilebilir.

Hizmet sağlayıcı adayı olarak *x* eşi, 4. adımda *y* eşinden gelen kaynak edinme talebi üzerine *y* eşinin arşivcisi olan A_y eşinden, eş tutarlılığı değerini ($PC(y)$) sorgular. Eğer *y* eşinin eş tutarlılığı değeri belirlenen bir eşik değerinin üzerinde ise *x* eşi verebileceği bant genişliği büyüklüğüne karar vererek bunu 5. aşamada döner. Kapsamlı deneyler sonucu $PC(y)$ için eşiği 0.5 olarak belirledik. Eğer $PC(y) > 0.5$ ise, *x* eşi hizmet sağlayıcı olarak *y* eşine bir bant genişliği sözü verir. Bant genişliği vaat ederken, *x* hem hizmet talep eden *y* eşinin tutarlılığını hem de devam eden hizmetlerini göz önünde bulundurarak şu şekilde hesaplar:

$$pBW_y = \frac{PC(y)}{tPC_x} \times tBW_x \quad (6)$$

$$tPC_x = \sum_{i \in R_x} PC(i) \quad (7)$$

Bu hesaplamalarda tBW_x değeri *x* eşinin toplam iletişi kanalı genişliğini (*bandwidth*), pBW_y değeri *y* eşine önerilen iletişim kanalı genişliğini ve tPC_x ise *x* eşinden hizmet alan, R_x ile gösterilen eşlerin eş tutarlılığı değerleri toplamını ifade etmektedir. Bir başka deyişle, *x* toplam iletişim kanalı genişliğini kendisinden hizmet talep eden eşlere, eş tutarlılığı değerlerine göre adil bir şekilde paylaştırmaktadır. Bu yaklaşım, eşler arası ağlarda güvenilir olmayı teşvik ederek, güvenilir bir eşin daha kaliteli hizmet almasını sağlayarak sistemin hizmet kalitesini (*Quality of Service*) yukarı taşımaktadır. Çünkü her bir eş, daha güvenilir eşlere daha fazla iletişim kanalı genişliği sağlamaktadır. Burada daha güvenilir olma durumunu sağlayan şey, alınan hizmete doğru geri bildirimde bulunmaktır ve eş tutarlılığı gösterge olarak değerlendirilmiştir.

Anlatımı kolaylaştırmak için **Hata! Başvuru kaynağı bulunamadı.**'de *y* eşi yalnızca *x* e şinden hizmet talep ediyor. Oysa ki gerçek uygulamada eşik değerinin üstünde güven

değerine sahip eşlerden topladığı bant genişliği önerilerinden en yüksek olanı seçmektedir.

Etkileşimin nasıl sona erdiğini **Hata! Başvuru kaynağı bulunamadı.**'de Etkileşim S onlandırılması etiketinden itibaren adımlar olarak gösterdik. Bir etkileşim sona erdiğinde ya da sonlandırıldığında, hizmet alan y eşi aldığı hizmete dair memnuniyet değerini $(s_i(x,y))$, hizmeti aldığı x eşinin bilgilerini tutan arşivcilerine gönderir. İkinci adımda A_x aldığı geri bildirim üzerinden geri bildirim tutarlılığı $(FC_i(x,y))$ ile güven değerini $T_i(x))$ hesaplayarak kaydeder. Üçüncü adımda A_x hesapladığı geri bildirim tutarlılığı $FC_i(x,y)$ değerini y eşinin arşivcisi A_y'ye gönderir. Son adımda ise A_y, y eşinin eş tutarlılığı değeri $PC(y)$'yi yeniden hesaplayarak kaydeder. Eğer y eşi geri bildiriminde yanlış bilgilendirme yaparsa y eşinin verdiği geri bildirim tutarlılığı $FC_i(x,y)$ düşük olacak, sonuçta y eşinin eş tutarlılığı değeri $PC(y)$ düşecektir.

4.6. Saldırı Modelleri

Bu çalışma kapsamında; bireysel saldırılar gerçekleştiren ve grup halinde hareket eden saldırgan türlerinin, sürekli toy saldırma (*naive)*, belli bir oranda iki yüzlü saldırma (*hypocritical*) ve belli güven aralığında, uyarlanabilir saldırma (*adaptive)* örüntülerini sergileyebildiği saldırgan/saldırı davranışları ele alınmıştır. Saldırgan modelleri ile saldırı örüntüleri birleştirilerek farklı saldırı etkinlikleri elde edilmiş, deney sonuçları bu çeşitlilik ile elde edilmiştir (Örneğin bireysel saldıran bir saldırgan modelinin iki yüzlü saldırı örüntüsü göstermesi ya da bu örüntüyü grup olarak sergilemeleri ayrı ayrı sınanmıştır).

4.6.1. Saldırgan Modelleri

Bireysel Kötücül Eşler: Diğer eşler ile herhangi bir iş birliğinde bulunmadan saldırılar gerçekleştiren eşlerdir.

İşbirlikçi Kötücül Eşler: Ağ içinde bulunan başka eşler ile birlikte hareket ederek birbirlerini kollayan, diğer eşlere ise saldıran eşlerdir. Birbirlerine yüksek geri bildirimler vererek ve saldırmayarak diğer geri bildirimlerin tutarlılığını da düşürmeye çalışırlar.

4.6.2. Saldırı Örüntüleri

Toy (Naïve) Saldırganlar: Aldığı tüm hizmetlerde geri bildirimleri, verdiği hizmetlerde ise sağladığı kaynağı bozarak ağı zehirleyen saldırganlardır.

İki yüzlü (Hypocritical) Saldırganlar: Aldığı ve verdiği hizmetlerde belli bir oranda saldıran eşlerdir. Örneğin saldırı oranı 0.2 seçilmişse, eş her 100 etkileşiminde 20 saldırı gerçekleştirir. Geri bildirim saldırıları ile iş birlikleri ile birlikte uygulandığında, tespiti en güç saldırı türlerinden birisidir.

4.7. Tutarlılık Temelli Modele Ait Deney Sonuçları

Tutarlılık temelli modelimizi 4 farklı senaryoda hizmet saldırıları, 4 farklı senaryoda ise geri bildirim saldırıları için sınadık. Söz konusu senaryoları Çizelge 4.1'de görebilirsiniz. Senaryoların her biri sırası ile; herhangi bir güven modelinin olmadığı ağ ortamında, tutarlılık temelli modelimizin etkinleştirildiği ağ ortamında ve en bilinir güven modellerinden olan EigenTrust modelinin etkinleştirildiği bir ağ ortamında sınanmış ve sonuçları alınmıştır. Herhangi bir sınıflandırmanın yapılmadığı bu deney ortamında, yalnızca etkileşimin saldırı olup olmadığı etiketlenmiş, bu etiketler üzerinden deney sonucunda saldırı oranlarının karşılaştırılması hedeflenmiştir.

Çizelge 4.1 Tutarlılık temelli model için saldırı senaryoları

Kötücül Davranış	**Kötücül eş oranı**	**İş birliği stratejisi**	**Hizmet / Geribildirim**
İki yüzlü saldırgan (0.2 saldırı olasılığı ile)	%20	İşbirlikçi veya Bireysel	Hizmet
Toy saldırgan	%20	İşbirlikçi veya Bireysel	Hizmet
İki yüzlü saldırgan (0.2 saldırı olasılığı ile)	%20	İşbirlikçi veya Bireysel	Geribildirim
Toy saldırgan	%20	İşbirlikçi veya Bireysel	Geribildirim

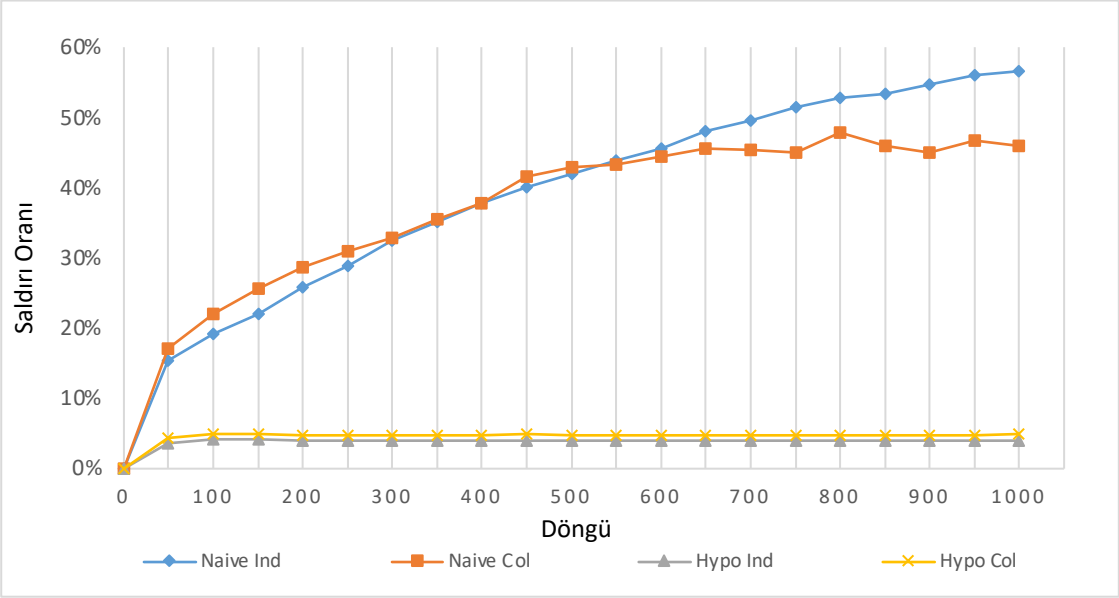
Bu tez çalışması için önerilen modelin değerlendirilebilmesi için *Peersim* [5] ortamı üzerinde benzetim modeli gerçekleştirilmiştir. *Peersim* döngü tabanlı (cycle-based) ve olay tabanlı (event-based) benzetim yeteneğine sahiptir. Bu çalışma kapsamında gerçek dünyadaki bir bilgi/belge temini sunan bir ağın döngü temelli olarak benzetimi yapılmıştır. Ancak önerilen modelin bu kullanım biçimi ile kısıtlanmasına gerek yoktur.

Tutarlılık temelli modelin sınanması için koşulan benzetimde ağda 10000 eş aynı anda etkin olarak yapılandırılmış, toplam döngü sayısı 1000 olarak uygulanmıştır. Her bir döngünün başlangıcında, eşler bir etkileşim başlatabilir (örn: kütük indirme), tamamlanan etkileşimleri sonuçlandırır ya da devam eden etkileşimleri ilerletebilir. Karşılaştırmalarda yer verilen Eigentrust için önemli bir parametre olan ön tanımlı güvenilir eş (pretrusted peer) sayısı 30 olarak belirlenmiştir. Ayrıca Eigentrust modelinde geri bildirim saldırıları ile mücadeleye dair bir öneri yer almadığı için yalnızca hizmet saldırılarında karşılaştırma yapılmıştır.

4.7.1. Hizmet Saldırıları Sonuçları

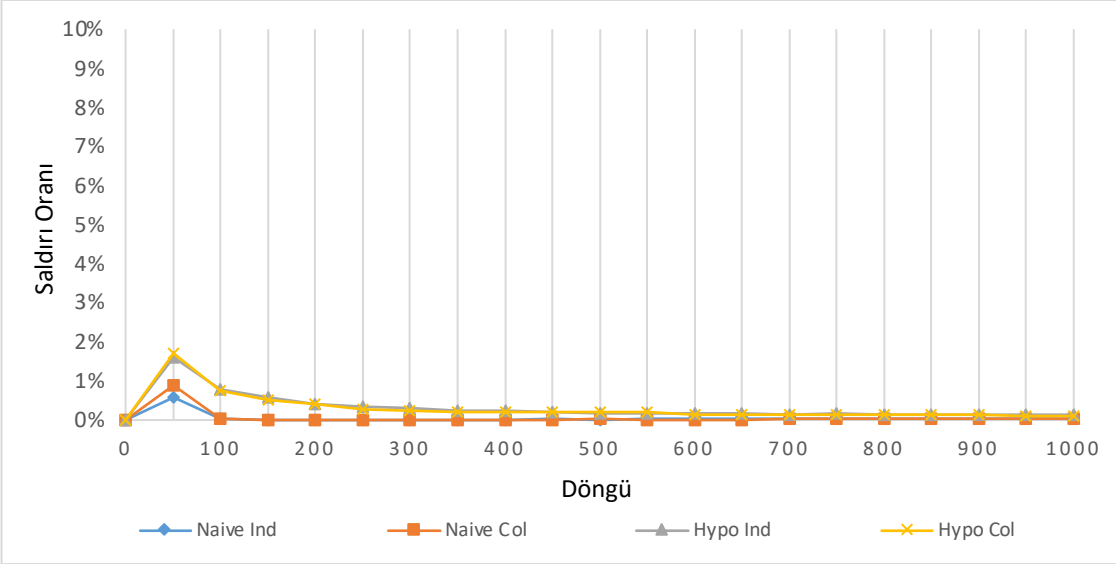
Bu başlık altında güven mekanizmasının olmadığı, tutarlılık temelli modelin etkinleştirildiği ve Eigentrust modelinin etkinleştirildiği deneyler ayrı ayrı verilmiştir. Şekil 4.2'de güven mekanizmasının olmadığı deney sonuçlarında görüldüğü gibi, iki yüzlü saldırganların olduğu senaryolarda, bireysel ya da işbirlikçi olması fark etmeksizin

%4.5 ile %5 arasında saldırı oranı gözlemlenmiştir. %20 saldırgan oranı ve 0.2 olasılıkla saldırı örüntüsünün doğal bir sonucu olarak %4 civarı saldırı ölçülmüştür. Ancak ilginç bir şekilde, %20 saldırgan oranının olduğu Naive saldırgan senaryolarında, saldırı oranının yavaş yavaş artarak %50'nin bile üzerine çıkabildiği gözlemlenmiştir. Saldırganların sistemdeki geçirdiği süre arttıkça daha fazla kaynağı biriktirerek bunları hizmet olarak sunduğunu gözlemlediğimizi aktarmamız uygun olacaktır.



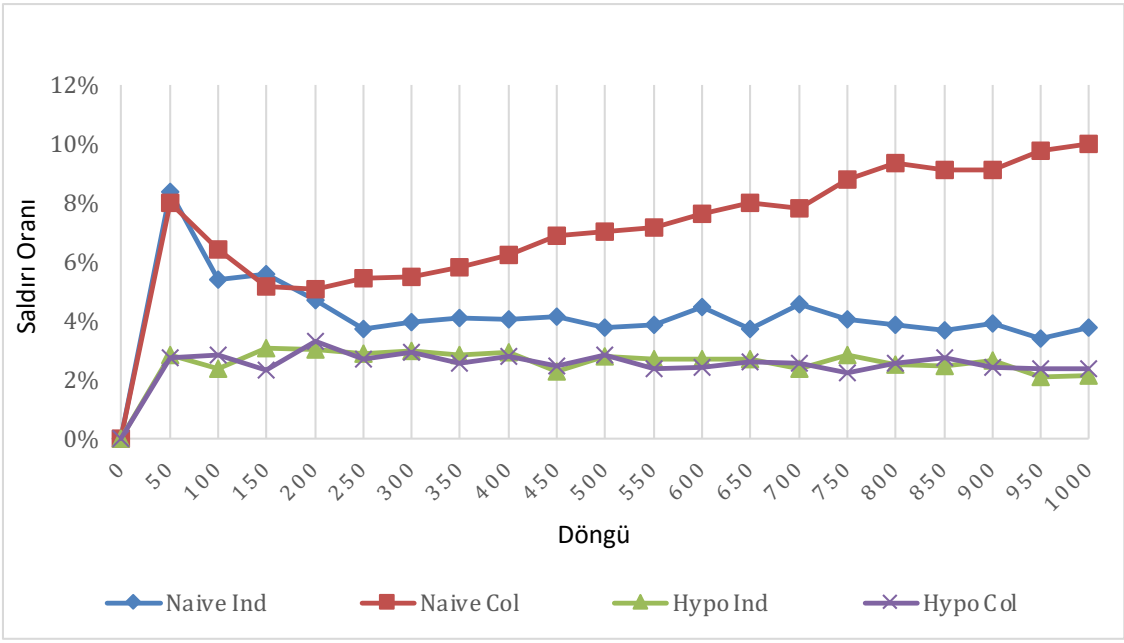
Şekil 4.2 Herhangi bir güven mekanizmasının olmadığı durumda farklı senaryolardaki hizmet saldırısı oranları

Şekil 4.3'de tutarlılık temelli modelin etkin olduğu eşler arası ağ ortamında ölçülen saldırı oranlarını görebilirsiniz. Deneyin ilk aşamalarında modelin kötücül eşleri tespit ettiğini ve 100. döngüden itibaren saldırı oranının hızla düşerek %0,5'in altına gerilediğini gözlemledik. 900.-1000. arası döngüler göz önüne alınarak güven mekanizmasının işletilmediği deneyler ile karşılaştırıldığında toy saldırgan senaryosu için başarılı indirme oranı %45'ten %99,5'e yükselmektedir. İki yüzlü saldırgan senaryolarındaki saldırı oranı beklentisi %4 civarında olduğu için söz konusu senaryolarda başarılı indirme oranında daha kısıtlı bir iyileşme alanı bulunmaktadır. İki yüzlü saldırgan senaryolarında başarılı indirme oranının %96'dan %99,5'a yükseldiği gözlemlenmiştir. Sonuç olarak tutarlılık temelli modelin etkin olduğu eşler arası ortamda başarılı etkileşim oranlarının 50. döngüden itibaren başarılı bir şekilde yükseldiği tespit edilmiştir.



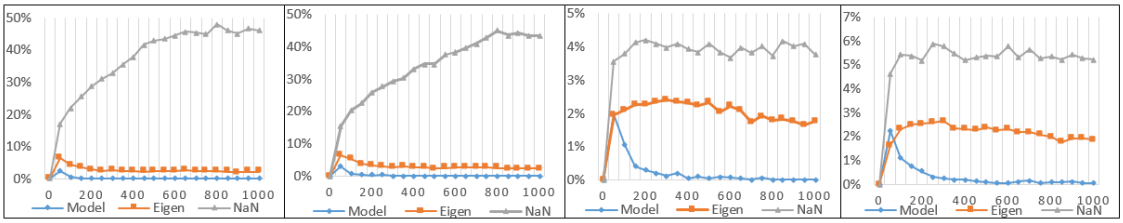
Şekil 4.3 Tutarlılık temelli istatistiksel modelin etkinleştirildiği durumda farklı senaryolardaki hizmet saldırısı oranları

EigenTrust modelinin hizmet saldırı senaryoları için elde ettiği başarımı **Hata! Başvuru k aynağı bulunamadı.**'te görebiliriz. Saldırı oranlarının nadiren %2'nin altına düştüğü gözlemlerimizde, iş birliği yapan toy saldırganların zamanla etkinliğini yavaş yavaş artırarak, deneyin son bölümünde %10'un üzerine çıktığını gördük.



Şekil 4.4 Eigentrust modelinin etkinleştirildiği durumda farklı senaryolardaki hizmet saldırısı oranları

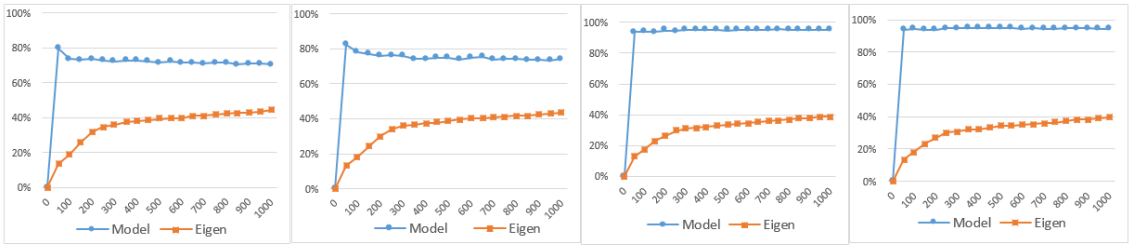
Son olarak, Şekil 4.5'te dört farklı senaryoda güven yönetiminin yapılmadığı, tutarlılık temelli model ile yapıldığı ve Eigentrust modeli ile yapıldığı durumlardaki sonuçları görülmektedir. Sonuçlar açık bir şekilde bir güven yönetimi ihtiyacını ortaya koyarken, tutarlılık temelli modelin saldırıları Eigentrust'tan çok daha iyi başarım ile engellediğini söyleyebiliriz.



a. Naive individual **b. Naive collaborative** **c. Hypocritical individual** **d. Hypocritical collaborative**

Şekil 4.5 Servis saldırılarında karşılaştırmalı özet

Bir güven yönetimi mekanizması için, sistemi koruma işini etkileşim sayısını azaltarak değil doğru yönlendirerek yapabilmesi önemli bir özelliktir. Bir başka deyişle, yanlış pozitif sayısını artırmadan doğru pozitif oranını artırmak daha kaliteli bir ağ ortamı sağlayacaktır. Bu amaçla benzetim ortamında topladığımız bir istatistik de, bir eşin bir hizmet talebi oluşturması ile bu hizmet talebinin karşılanması oranıydı. Yani, bir eş hizmet talep ettiğinde, saldırmayacak bir eşten bu hizmeti alabilmesini sağlamak, sistemin hizmet kalitesini de artıran bir ölçüt olarak görülebilir. Şekil 4.6'da Eigentrust ile tutarlılık temelli modelin başarılı indirme girişimi oranlarını görebiliriz. Tutarlılık temelli modelin farklı senaryolar için %70-90 arasında indirme girişimini başarı ile tamamladığı ölçülürken, Eigentrust'ın sistemi korumak için girişimleri engellediği, %40'ın altında bir başarılı indirme girişimi oranı sağladığı gözlemlenmiştir. Tutarlılık temelli model, %20 toy saldırganın hizmet almasını engelleyen bir sistem sunduğu için, indirme girişimlerinin oranının %80 bölgesinde olması iyi bir işaret olarak değerlendirilebilir. Benzeri şekilde, iki yüzlü saldırgan senaryolarında indirme girişimlerinin %95'i tutarlılık temelli modelde başarılı bir şekilde hizmet alabilirken, Eigentrust'ta bu oran yine %40'ın altında kalmaktadır.

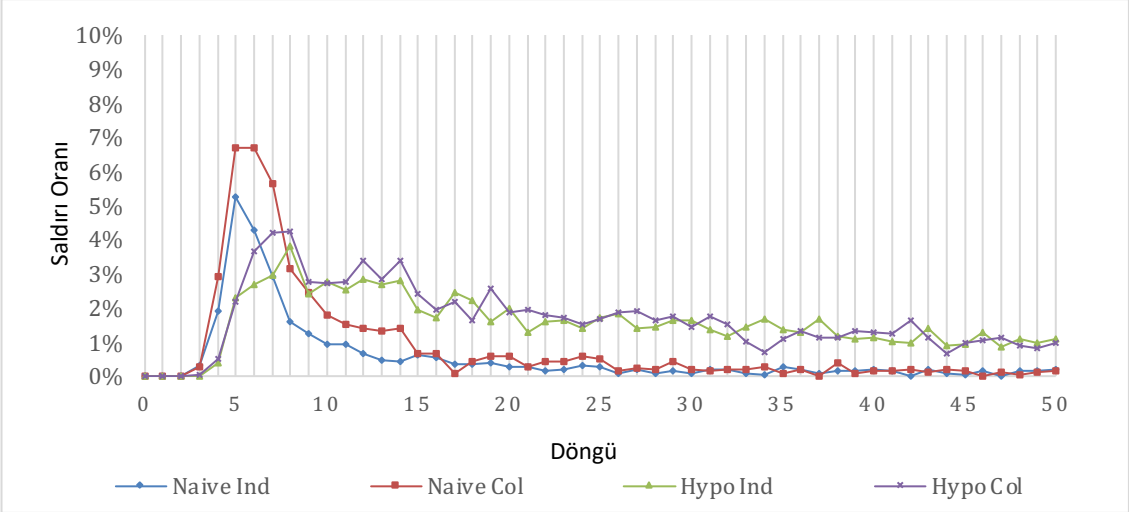


a. Naive individual b. Naive collaborative c. Hypocritical individual d. Hypocritical collaborative

Şekil 4.6 İndirme girişimi ile indirme sayısı arasındaki oran üzerinden Eigentrust ile tutarlılık temelli modelin karşılaştırılması

4.7.2. Önyükleme (Bootstrap) Süreci Sorunu

Hizmet saldırılarında deney sonuçları gösterdi ki, tutarlılık temelli model sınandığı senaryolarda yüksek bir başarım sergilemektedir. Ancak dikkat çekici bir nokta olarak, deneysel gözlemlerin ilk bölümünde, bir süre saldırı oranları daha yüksek çıkmaktadır. Buradaki sonuçlara yakın plandan bakıldığında daha iyi anlaşılacaktır. Şekil 4.7'de her bir senaryo için ilk 50 döngüde hangi oranda saldırı gerçekleştiğine yakından bakıldığında, özellikle ilk 5-10 benzetim döngüsünde saldırı oranları en yüksek seviyelere çıkmaktadır. 20. benzetim döngüsünden itibaren ise toy saldırgan senaryolarında saldırı oranı %0,5 değerinin altına kadar hızla gerilerken, iki yüzlü saldırgan senaryolarında %2 civarında kalmaktadır.

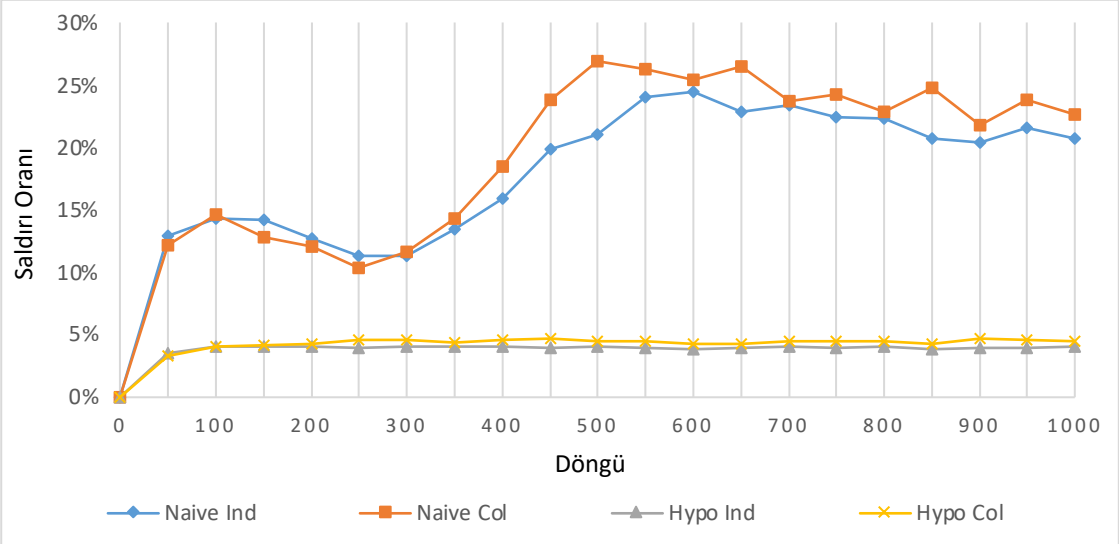


Şekil 4.7 Önyükleme sürecinde tutarlılık temelli modelin etkin olduğu ortamdaki ilk 50 döngünün saldırı oranları

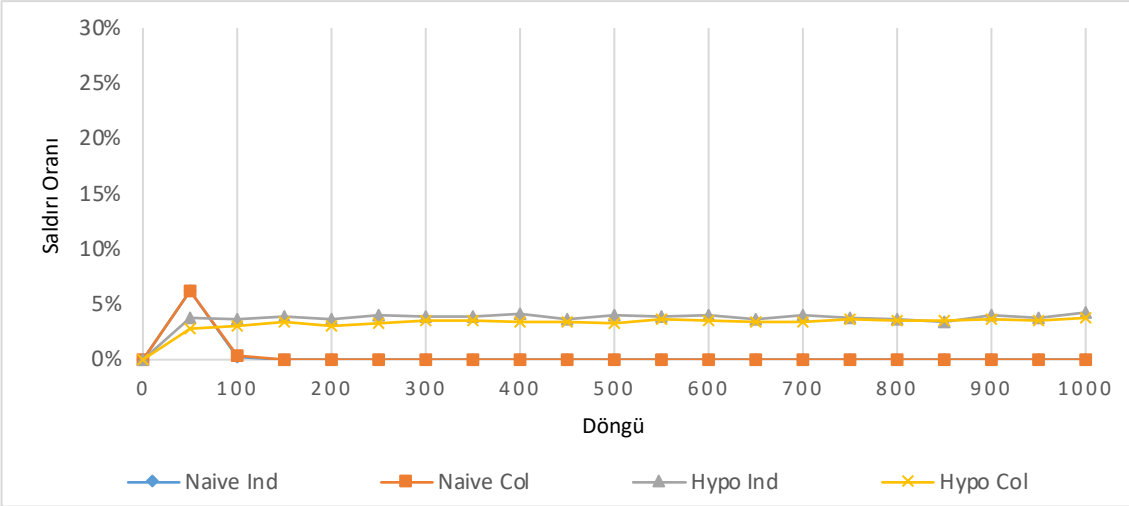
4.7.3. Geribildirim saldırıları sonuçları

Modelimizi daha ileri seviyede değerlendirebilmek için, tutarlılık ile geri bildirim arasındaki ilişkiyi çözümlemeye çalıştık. Hizmet saldırılarında karşılaştırmalarda yer verdiğimiz EigenTrust modeline, geribildirim saldırıları kapsamında bir çözüm sunmadığı için bu kesimdeki karşılaştırmalarda yer verilememiştir.

Herhangi bir güven mekanizmasının olmadığı bir ortamda saldırı senaryolarını incelediğimizde, Şekil 4.8'da görüleceği üzere, toy saldırgan senaryolarında saldırı oranları %30'a kadar yükselmiş ve %20-30 arasında değerler gözlemlenmiştir. İki yüzlü saldırgan senaryolarında ise, geri bildirim saldırıları %5 civarında durağan bir görünüm sergilemektedir.



Şekil 4.8 Herhangi bir güven mekanizmasının olmadığı durumda farklı senaryolardaki geri bildirim saldırısı oranları



Şekil 4.9 Tutarlılık temelli istatistiksel modelin etkinleştirildiği durumda farklı senaryolardaki geri bildirim saldırısı oranları

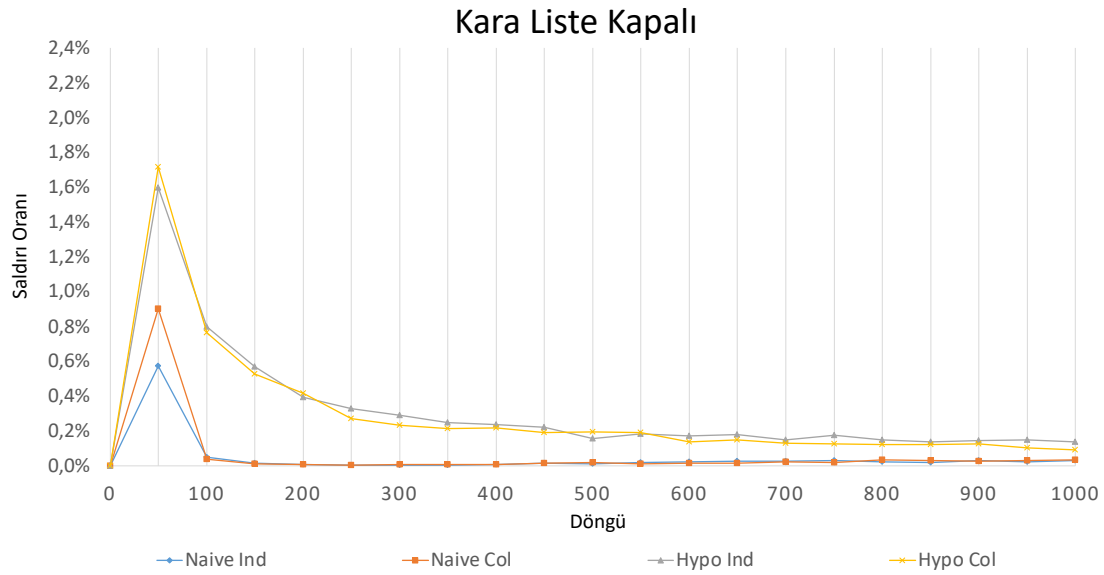
Tutarlılık temelli istatistiksel modelin etkinleştirildiği durumda elde edilen sonuçları Şekil 4.9'den inceleyebiliriz. toy saldırganlar tespit edilerek erken aşamalardan itibaren sistemdeki etkinlikleri ortadan kaldırılmış görünmektedir (toy işbirlikçi saldırganlar ile toy bireysel saldırganların sonuçları çok benzer olduğu için üst üste çakışmaktadır). İki yüzlü saldırgan senaryolarında, geri bildirim saldırı oranının %3,5 ile %4 arasında durağan kaldığını gözlemledik. Bu açıdan değerlendirdiğimizde, hizmet saldırılarını tespit etmekte çok başarılı olduğunu gözlemlediğimiz tutarlılık temelli istatistiksel modelimizin, iki yüzlü geri bildirim saldırganlarını tespit etmekte aynı derecede başarım sergileyemediğini söyleyebiliriz. Buradaki başarımın aynı derecede iyi olmamasının nedeninin, kötücül geribildirimlerin %100 doğrulukla tespit edilmesinin mümkün olmamasından kaynaklı olduğu değerlendirilmiştir.

4.8. Tutarlılığa Dayalı Model Çözümleme ve İyileştirme Çalışmaları

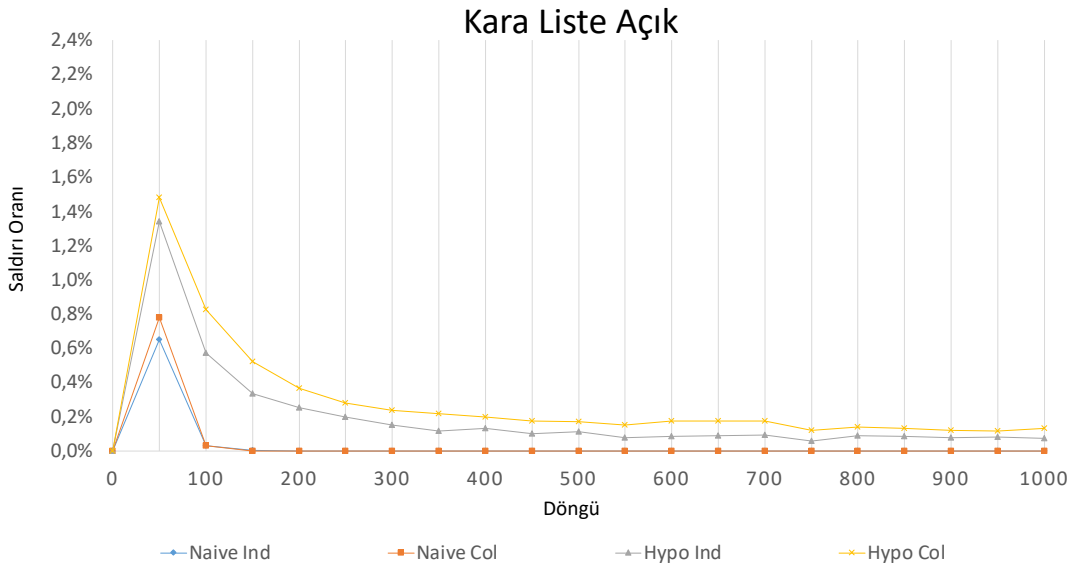
4.8.1. Kara liste iyileştirmesi

Kara liste (blacklist), bir eş ile ilgili elde edilen olumsuz bir deneyime dayanarak herhangi olası bir etkileşim için söz konusu eşin yasaklanmasını sağlayan listedir. Literatürde eşlerin kendileri için kara liste tutulduğu [66] ya da bu kara listeleri paylaşarak diğer eşleri etkileyecek şekilde itibara yansıtıldığı uygulamalar [67] mevcuttur. Biz çalışmamızda

kara liste uygulamasını yeni bir saldırı türüne konu olmaması için, eşlerin kendi yerel listelerini tutmaları şeklinde sınırlandırıp etkilerini inceledik.



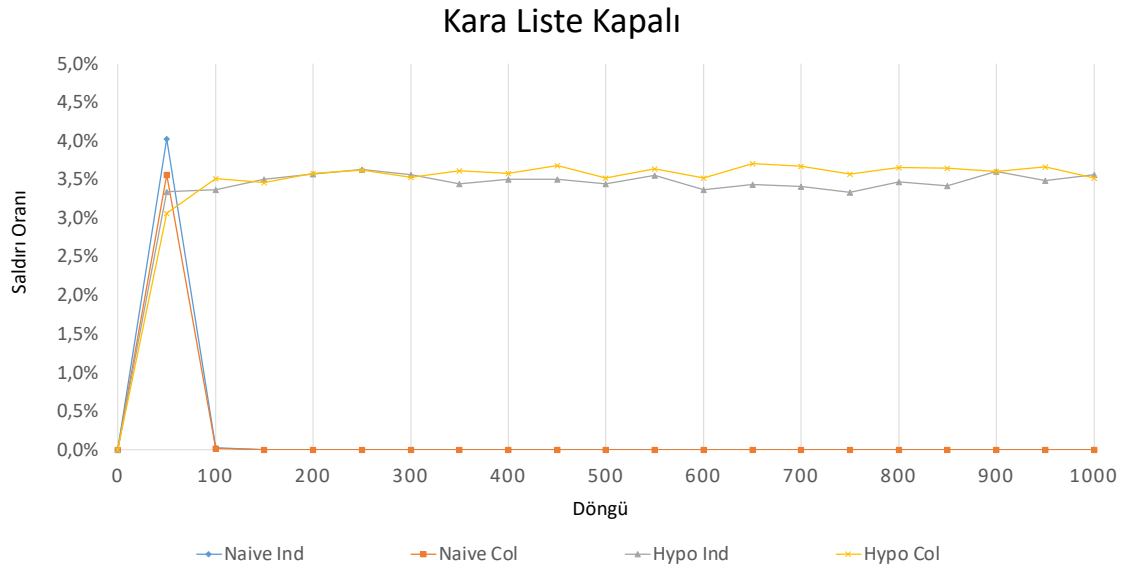
Şekil 4.10 Kara liste kapalı iken hizmet saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları



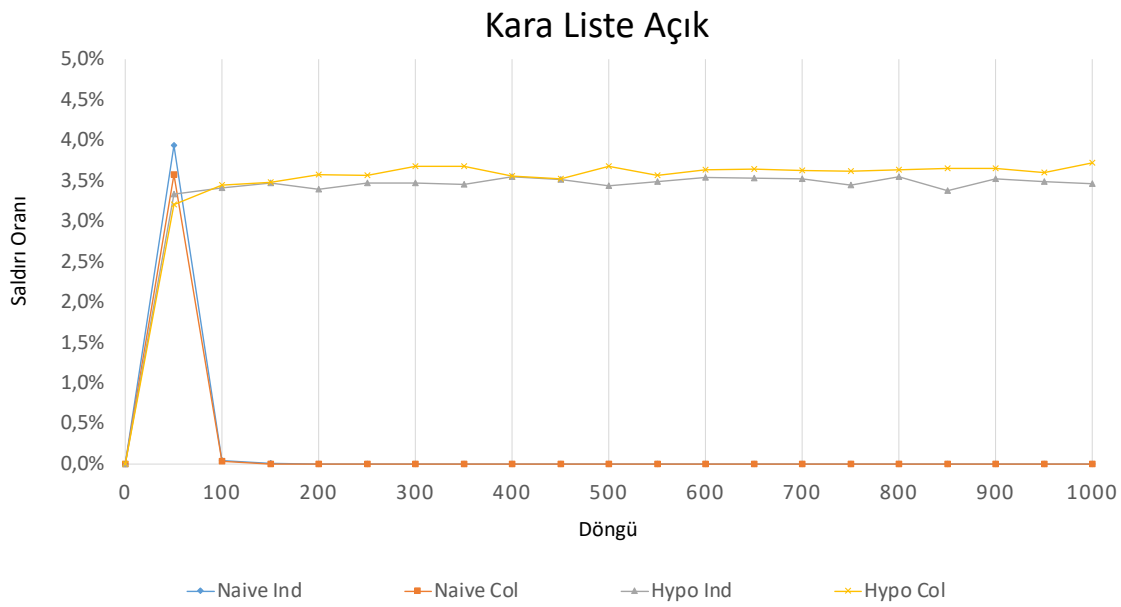
Şekil 4.11 Kara liste etkinleştirildiğinde hizmet saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları

Şekil 4.10'de verilen kara listenin etkin olmadığı deney sonuçları ile Şekil 4.11 Kara liste etkinleştirildiğinde hizmet saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları'de raporlanan, kara listenin etkinleştirildiği deney sonuçlarını karşılaştırdığımızda; en yüksek saldırı oranına ulaşan işbirlikçi iki yüzlü saldırgan senaryosunu, bu tepe noktası için %10'un üzerinde iyileştirmiştir. Ayrıca iki yüzlü saldırılarının oranını ilerleyen döngü benzetimi adımlarında %20'nin üzerinde iyileştirmektedir. toy saldırgan senaryolarında ise her ne kadar 100. döngüden itibaren 0'a yakınsamış olsa da, başlangıç aşamasında %10 civarı biri iyileşme sağladığı gözlemlenmiştir.

Ayrıca, geri bildirim saldırıları için de kara listenin etkilerini araştırdık. Sırası ile kara listenin kapalı ve açık olduğu deney sonuçlarını gösteren Şekil 4.12 ve Şekil 4.13'ü incelediğimizde, kara listenin geri bildirim saldırılarına olumlu ya da olumsuz bir etkisinin olmadığını söyleyebiliriz.



Şekil 4.12 Kara liste kapalı iken geri bildirim saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları



Şekil 4.13 Kara liste etkinleştirildiğinde geri bildirim saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları

4.8.2. Medyan iyileştirmeleri

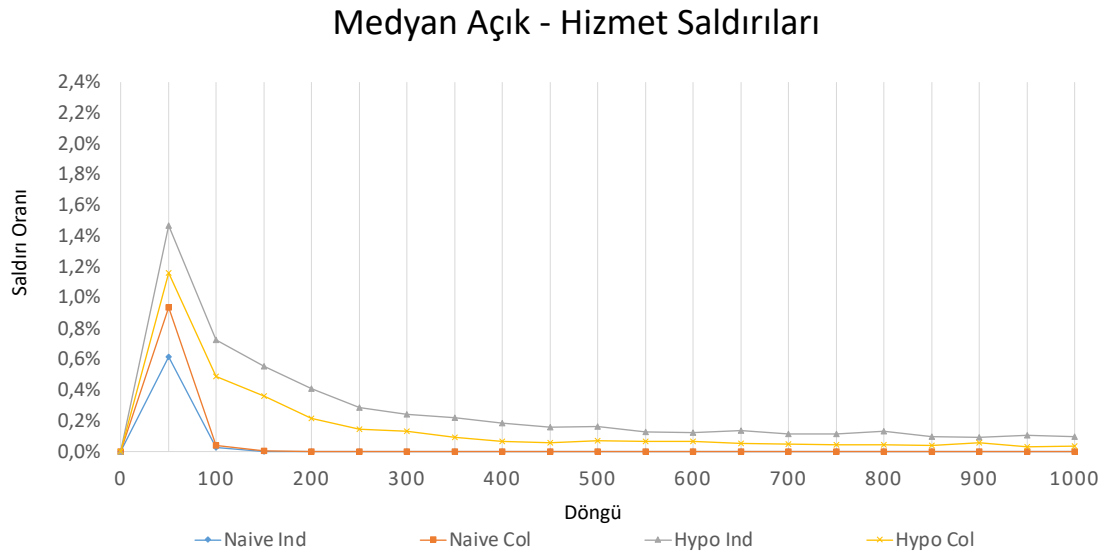
Medyan (ortanca), bir veri serisini küçükten büyüğe sıralandığında, seriyi ikiye ayıran elemana verilen isimdir. Genelde ortalama yerine kullanılan bir istatistiktir. Çalışmamızda güven değerini hesapladığımız Eşitlik 4'ü farklı medyan ve ortalama ile iyileştirmeye çalıştık.

$$T_{M_i}(x) = \propto . M_i\ (x,y) + (1-\propto) . T_{(M_{i-1})}\ (x) \quad (8)$$

$$M(x,y) = (M\{E_j\ (x,y)_{j=n-2}^{n}\}) \quad (9)$$

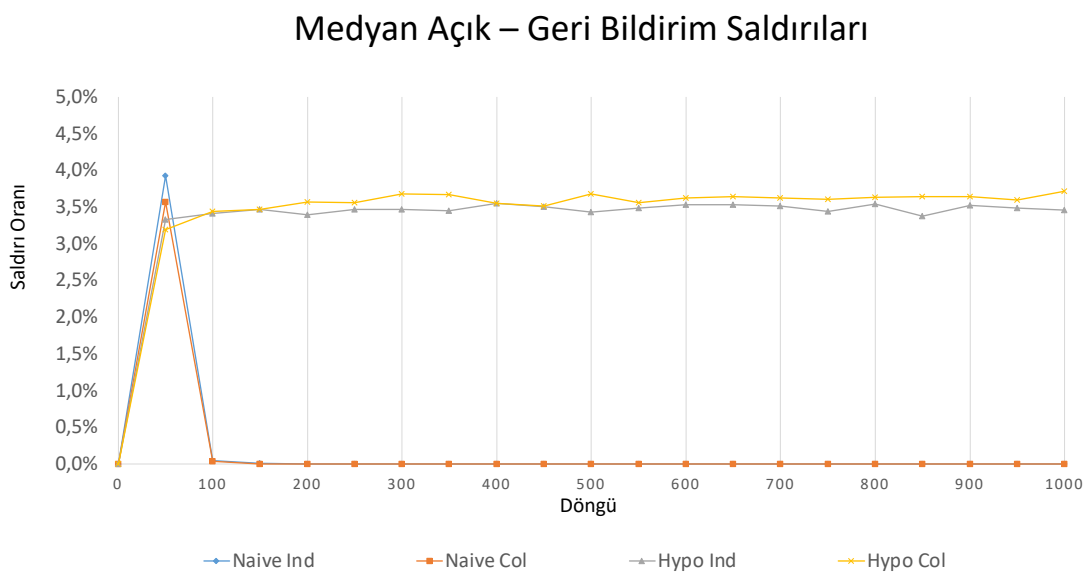
Eşitlik 8'de yer alan $M_i\ (x,y)$, Eşitlik 9'da görülebileceği gibi, son 3 geri bildirimin Eşitlik 5 ile elde edilmiş skorlarının medyan değeridir. 3 değeri, pencere boyu olarak benzetim ortamında yapılandırılmıştır. Yapılan kapsamlı deneyler sonucunda 3, 5 ve 10 sınanmış, en iyi sonuç veren pencere boyu yapılandırması 3 olarak belirlenmiştir. Benzeri bir uygulama eş tutarlılığının hesaplandığı Eşitlik 3 için de sınanmış, ancak sonuçlarda bir iyileşme olmadığı için modele eklenmemiştir.

Çalışma kapsamında Eşitlik 4 yerine Eşitlik 8'i kullanarak hem hizmet saldırılarını hem de geri bildirim saldırılarına etkisini gözlemledik. Şekil 4.14'de elde ettiğimiz hizmet saldırılarına dair sonuçları, Şekil 4.10'daki yalın uygulama ile kıyasladığımızda, işbirlikçi iki yüzlü saldırganların tepe noktasının %40 kadar iyileştiğini, öyle ki neredeyse toy işbirlikçi saldırganlar kadar iyi saldırı engellemeye başladığını gözlemledik. İlerleyen döngülerde ise tüm iki yüzlü saldırı senaryosu sonuçları daha hızlı bir şekilde 0'a yakınsadığı tespit edilmiştir.



Şekil 4.14 Medyan uygulandığında hizmet saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları

Medyan uygulamasının geri bildirim saldırılarının engellenmesi üzerindeki etkisini görmek için **Hata! Yer işareti başvurusu geçersiz.** ile yalın sonuçları veren Şekil 4.12 incelenebilir. Sonuçlar gösterdi ki, kara liste uygulamasında olduğu gibi medyan uygulaması da tek başına geri bildirim saldırılarını engelleme yeteneklerinde bir iyileşme sağlamamaktadır.



Şekil 4.15 Medyan uygulandığında geri bildirim saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları

4.8.3. Yerel güven iyileştirmesi

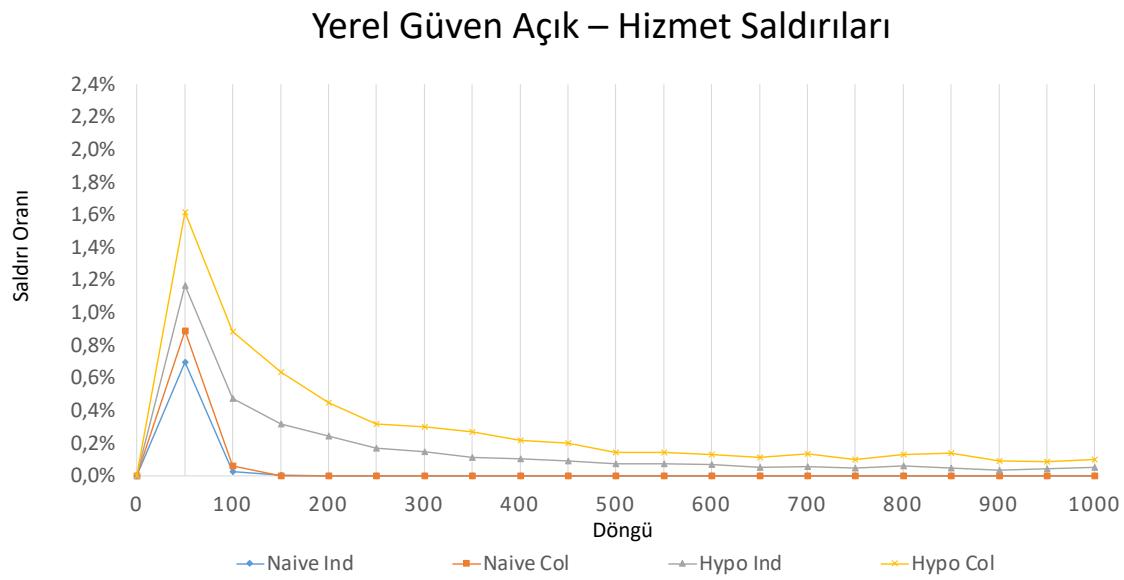
Çalışmamızın bu kesimine kadar güven hesaplamalarını yaparken tüm eşlerin aynı sonuca ulaşacağı aşamalardan geçtik. Herhangi bir x eşi için ağdaki herhangi bir y eşi güven değeri hesapladığında aynı sonuca varacaktır. Ancak güveni yalnızca itibar üzerinden değil, kişisel deneyimlerden de etkilenecek şekilde güncellemek için başvurduğumuz bir yöntem olarak yerel güven hesaplamaları, kara listeden sonra her eşin kendi içerisinde karar mekanizmalarına dahil ettiği ikinci bir girdi oldu. Literatürde kara liste gibi yerel güven değeri de çalışılmış bir yaklaşımdır [16,19].

$$L(x,y) = \frac{\sum s_i(x,y)}{[\mathbb{F}(x,y)]} \tag{10}$$

$$T(x,y) = T(y) * 0{,}5 + L(x,y) * 0{,}5 \tag{11}$$

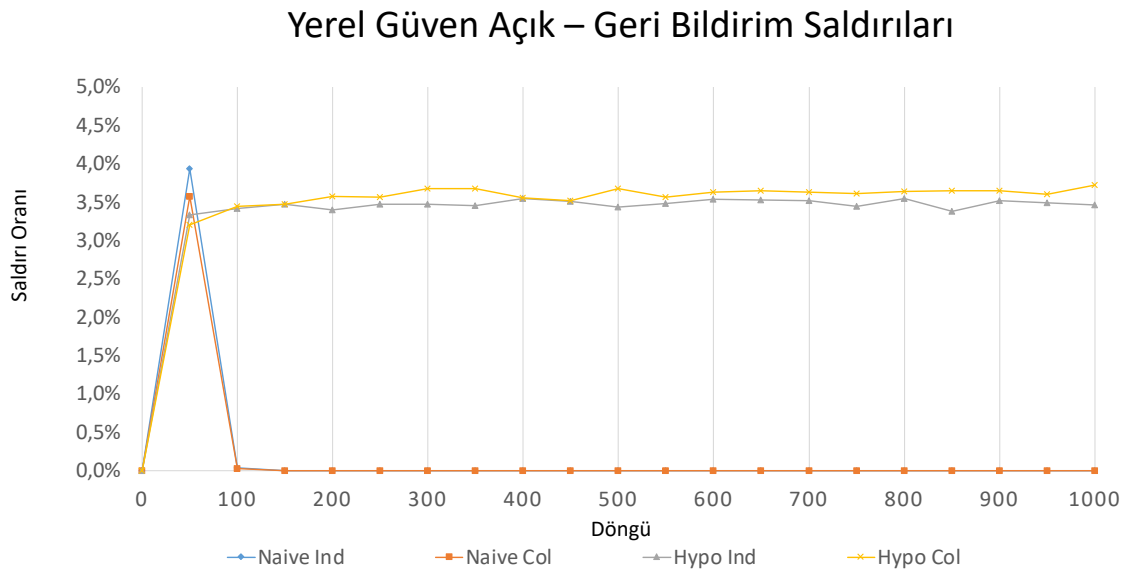
$L(x,y)$ x eşinin y eşi ile ilgili yerel güven değerini göstermektedir ve geçmiş etkileşimlerindeki memnuniyet değerlerinin ortalamasıdır. Eşitlik 11 ile karar verilme anında hesaplanan güven değeri, y eşinin Eşitlik 4 veya Eşitlik 8 ile elde edilen güven $(T(y))$ değerinin, Eşitlik 10 ile hesaplanan güven değeri ile birleştirilmesi ile elde

edilmiştir. Biz deneylerimizde genel güven değeri ile yerel güven değerini birleştirirken, yaptığımız deneyler sonucu her iki bileşenin için 0.5'in en iyi ağırlık olduğunu gördük ve tüm deneylerde 0.5'i aktif olarak kullandık. Ancak diğer çalışmalarda bu alanın esnek olarak kullanılabilmesi için benzetim ortamımızda yapılandırılabilir olarak yer verdik.



Şekil 4.16 Yerel güven hesaplaması uygulandığında hizmet saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları

Yerel güven hesaplamasını etkinleştirerek deneyleri yenilediğimizde, Şekil 4.16'de verilen sonuçlara göre, hizmet saldırılarına ait sonuçların bireysel saldırganlardan hem iki yüzlü hem toy saldırgan senaryolarının tepe noktasında, Şekil 4.10'daki sonuçlara göre %15 civarında iyileştiğini, işbirlikçi saldırganlar için herhangi bir iyileşme olmadığını gözlemledik. İlerleyen döngülerde özellikle iki yüzlü bireysel saldırı oranlarının 0'a çok daha hızlı ve etkili bir şekilde yaklaştığını gözlemledik. Ne var ki, bu yaklaşımın geri bildirim saldırılarını daha etkin bir şekilde engellediğini söyleyebileceğimiz bir deneysel sonuç elde edemedik. Şekil 4.17'de görüleceği gibi, yerel güven değeri kullanımı geri bildirim saldırıları üzerinde ciddi bir etki yapmamıştır.

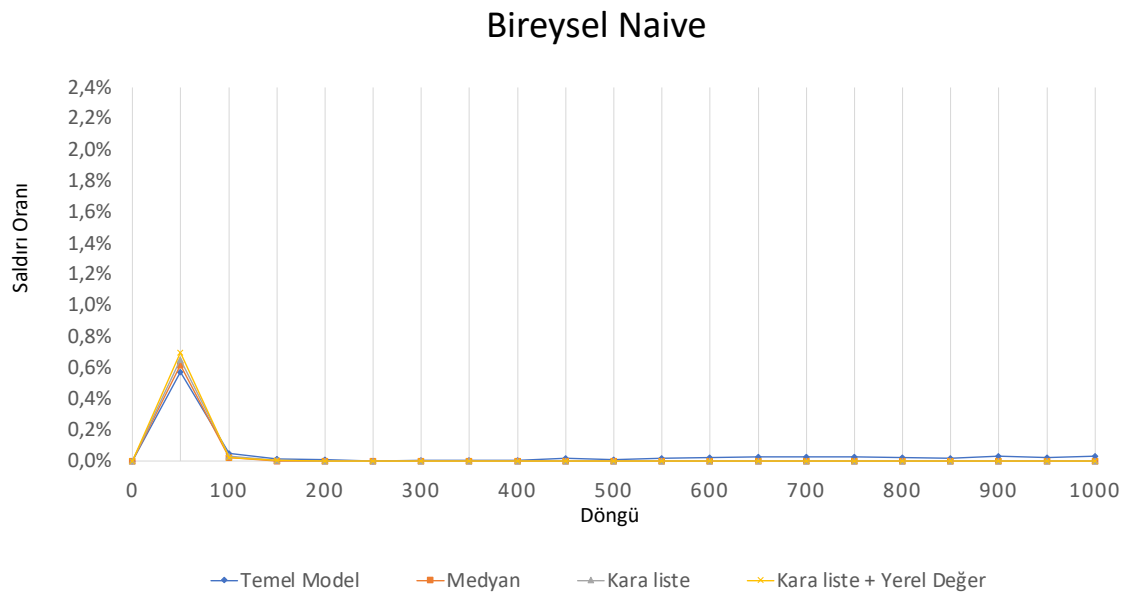


Şekil 4.17 Yerel güven hesaplaması uygulandığında geri bildirim saldırısı deneylerinde tutarlılık temelli model ile elde edilen saldırı oranları

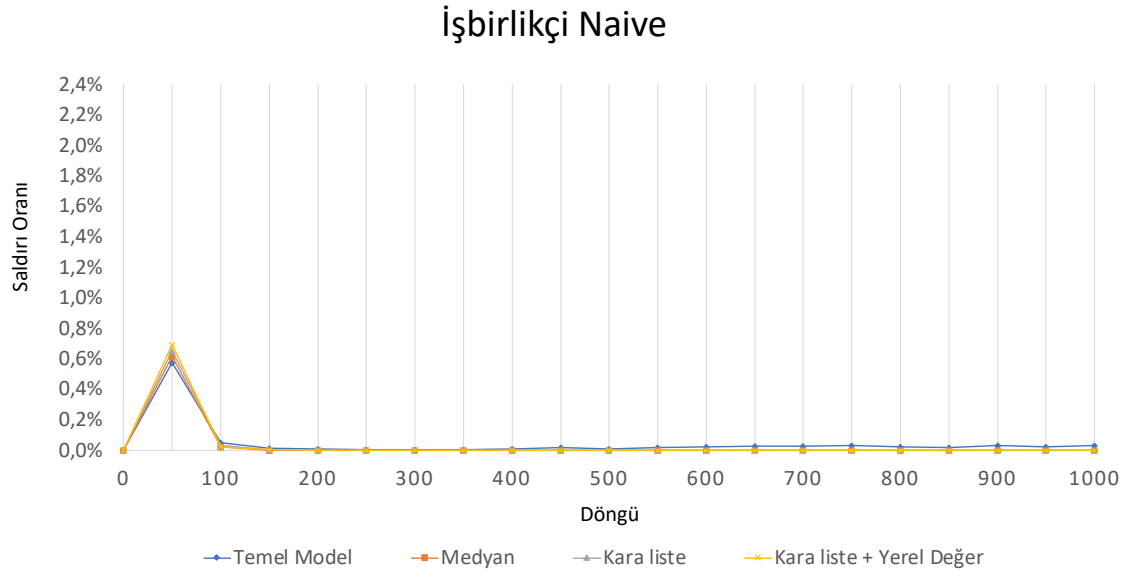
4.8.4. İyileştirmelere dair bazı çözümlemeler

Kara liste, medyan ve yerel güven hesaplamaları ile modelimizin saldırılara karşı başarımını artırmayı hedefledik ve her birinden farklı saldırı senaryolarında farklı düzeylerde olumlu sonuçlar aldık. Şekil 4.18, Şekil 4.19, Şekil 4.20 ve Şekil 4.21'de sırası ile bireysel toy, işbirlikçi toy, bireysel iki yüzlü ve işbirlikçi iki yüzlü saldırgan senaryolarının tutarlılık temelli model, medyan eklentili model, kara listeli eklentili model ve kişisel değerlendirmelerin birlikte olmasını sağlayan kara liste ile yerel güven hesaplamasını içeren modelin karşılaştırmalı sonuçları verilmiştir. Bireysel ve işbirlikçi toy saldırganlar tutarlılık temelli model ile çok yüksek bir oranda başarı ile engellenmekteydi. İyileşme alanının çok daralmasının da etkisi ile yapılan eklentilerin sonuçları birbirine oldukça yakın çıkmıştır. Şekil 4.20 ve Şekil 4.21 incelendiğinde, iki yüzlü saldırganların engellenmesinde modelde medyan uygulamasının genel olarak diğer senaryolara göre daha iyi sonuç verdiği gözlemlenmiştir. Ayrıca çözümlemesi yapılan tüm eklentilerin modeli iki yüzlü saldırılarına karşı farklı seviyelerde de olsa iyileştirdiğini söyleyebiliriz. Söz konusu model iyileştirmelerinin hiçbirinin geri bildirim saldırılarında herhangi anlamlı bir iyileşme sağlamadığını da gözlemledik. Yani özetle, söz konusu geliştirmeler hizmet saldırılarında iki yüzlü saldırganların

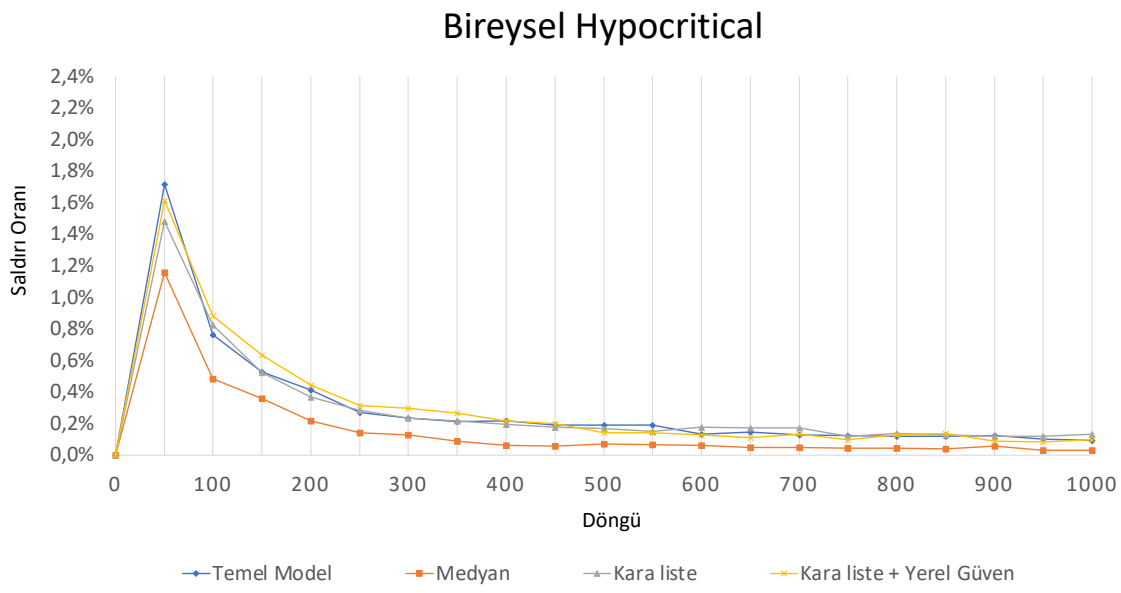
etkisizleştirilmesine karşı modele olumlu katkı vermiş, diğer senaryolar için etkisi çok kısıtlı kalmıştır.



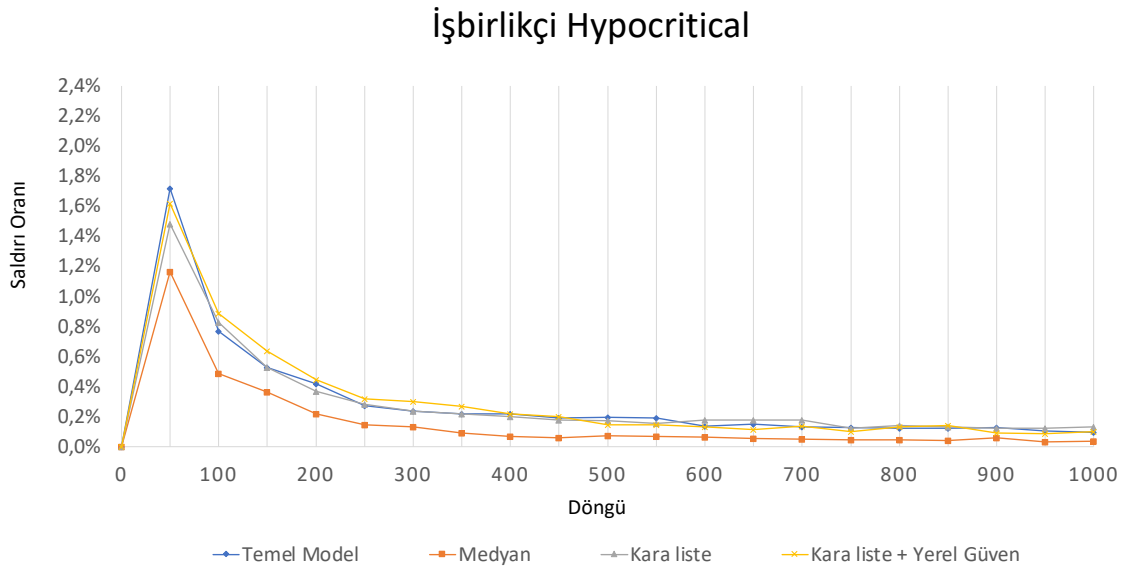
Şekil 4.18 Karşılaştırmalı bireysel toy hizmet saldırısı senaryosu sonuçları



Şekil 4.19 Karşılaştırmalı işbirlikçi toy hizmet saldırısı senaryosu sonuçları



Şekil 4.20 Karşılaştırmalı bireysel iki yüzlü hizmet saldırısı senaryosu sonuçları

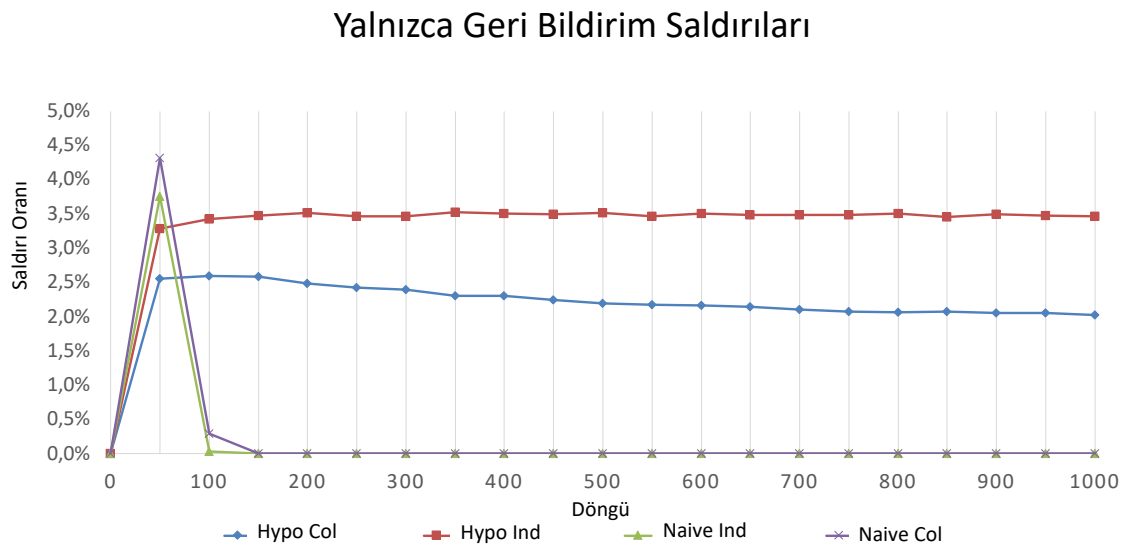


Şekil 4.21 Karşılaştırmalı işbirlikçi iki yüzlü hizmet saldırısı senaryosu sonuçları

4.8.5. Yalnızca Geri Bildirim Saldırı Senaryosu

Literatürde yaygın olan çalışma alanının hizmet saldırıları olduğuna değinmiştik. Bazı çalışmalar ise yalnızca geri bildirim saldırılarına odaklanmıştır [35]. Bu çalışma kapsamında biz esasen hem hizmet saldırılarına hem de geri bildirim saldırılarına odaklanmıştık ancak tutarlılık temelli modelimizin yalnızca geri bildirim saldırılarını

engellemedeki başarısını da görmek istedik ve bu konuda bir deney çalışması üzerinden çözümleme yaptık.



Şekil 4.22 Yalnızca geri bildirim saldırısı senaryosu sonuçları

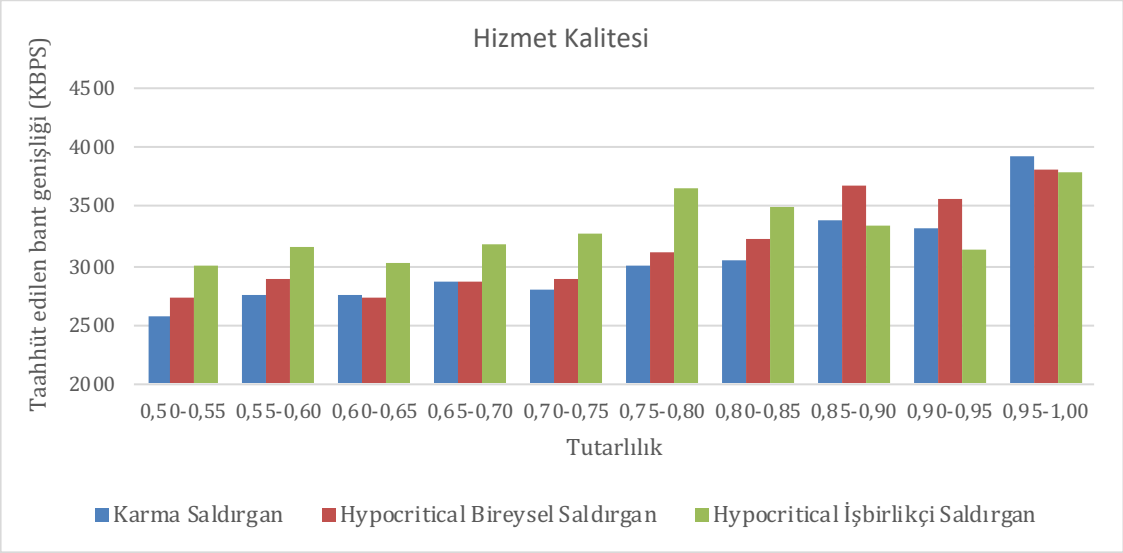
Yalnızca geri bildirim sonuçları Şekil 4.22'de verilmiştir. Beklentimiz yalnızca geri bildirim verirken kötücül davranılmasının modeli daha zor durumda bırakacağı ve bu alanda karşılaşacağımız potansiyel bir sorununun üzerine gitmemiz gerekeceği idi. Ancak sonuçlar bize gösterdi ki, yalnızca geri bildirim saldırısı gerçekleştirildiği senaryolar, Şekil 4.12'de görülebilecek normal geri bildirim saldırılarının olduğu senaryoya göre 100. Döngü adımından itibaren model için daha az zorlayıcı diyebileceğimiz bir başarım sergilemiştir. Bu sonuçlar ışığında yalnızca geri bildirim saldırılarını saldırgan senaryolarımız arasında yeniden konu etmemeyi tercih ettik.

4.9. Hizmet Kalitesi

Temel modelimizde kullandığımız etkileşim döngüsü, Eşitlik 6'ya göre tutarlı bir bant genişliği belirleme süreci kullanarak dürüstlüğü destekler. Bir eş ne kadar dürüst olursa, bant genişliği o kadar yüksek hizmet alarak hizmet kalitesini artırabilir. Bu şekilde dürüstlük, tutarlılık yoluyla teşvik edilir. Tutarlılık ve taahhüt edilen bant genişliği arasındaki ilişkiyi

Şekil 4.23'da görebiliriz. Bu kesimdeki deneylerin parametreleri Çizelge 5.6'daki değerlere göre ayarlanmıştır. Elde edilen sonuçlar incelendiğinde, modelin hizmet

kalitesini eşler arasında adil bir şekilde yönettiğini söyleyebiliriz. Daha az tutarlılığa sahip bir istek sahibi eş, potansiyel sağlayıcı eşler tarafından reddedildiğinden, şekilde 0.50'den daha küçük tutarlılık değerleri yoktur.



Şekil 4.23 Tutarlılık temelli modele ait hizmet kalitesi (QoS) sonuçları

5. MAKİNE ÖĞRENİMİ MODELİ

Tutarlılık temelli güven hesaplaması yöntemimiz ile elde ettiğimiz sonuçlar gösterdi ki temel aldığımız bu model iyi bir başarım ile saldırıları engelleyerek:

- *Toy* hizmet saldırıları başlangıç aşamasından sonra 0'a yaklaşıyor.
- *Toy* geri bildirim saldırıları başlangıç aşamasından sonra 0'a yaklaşıyor.
- *İki yüzlü* saldırganlarda belli bir seviyeye kadar hizmet saldırı oranları düşürülüyor, ancak bir noktadan sonra yakalanamıyor.
- *İki yüzlü* saldırganlarda geri bildirim saldırıları bir miktar azalırken, saldırıların önemli bir kısmı yakalanamıyor.

Yukarıdaki bazı saldırı türlerinde başarımı artırabilmek için temel aldığımız tutarlılığa dayalı modelimizi makine öğrenimi algoritmaları ile güçlendirmeyi hedefledik. Ayrıca literatürde gördük ki, güven yönetiminde makine öğreniminin kullanıldığı araştırmalar daha çok sosyal ağlara odaklanmış durumda ve makine öğrenmesinin çalışıldığı eşler arası ağlarda güven yönetimi araştırmalarında dikkate değer seviyelerde bir açık var. Bunun yanı sıra, eşler arası ağlar dışında kalan alanlar da dahil, güven yönetiminde makine öğrenmesine başvuran çalışmalar genel itibari ile yalnızca tek taraflı bir değerlendirmeyi kapsıyor. Biz bu çalışmamızda, hem hizmet verilmesi sırasında saldırıların sınıflandırmasına çözüm öneriyoruz, hem de hizmetin değerlendirildiği geri bildirim aşamasında ortaya çıkabilecek saldırıların sınıflandırılmasını bir çözüm önermeye çalışıyoruz.

5.1. Öznitelikler

Makine öğrenmesi çalışmalarında kullanılacak yöntemlerin başarımı büyük oranda özniteliklerin doğru seçimine ve bu özniteliklerin doğru konumlandırılmasına bağlıdır. Biz de bu çalışmamızın makine öğrenmesi aşamasında tutarlılık temelli modelimizden elde ettiğimiz özniteliklerin yanında, kural tabanlı *Prefix-Suffix* algoritmasını uygulayan Yahyaoui'nın [50] çalışmasında kullandığı öznitelikleri ve bilgi teorisindeki (*information theory*) entropi kavramından esinlenerek ürettiğimiz farklı öznitelikleri algoritmalarda kullanarak en iyi sonuçları elde etmeye çalıştık. Bu çalışma için yüzlerce öznitelik

üretilmiş, 60 civarı öznitelik çalışmaya dahil edilmiştir. Öznitelikleri 3 başlıkta açıkladık; (1) tutarlılık temelli modelden gelen öznitelikler, (2) diğer çalışmalardan ithal ettiğimiz ya da esinlendiğimiz öznitelikler, (3) ikinci faz çalışmalarımızda ürettiğimiz öznitelikler.

Tutarlılık temelli öznitelikler

Tutarlılık temelli istatistiksel modelin uygulaması sırasında toplanan istatistiksel bilgilerden elde edilen öznitelikler bu gruba dahildir. **Hata! Başvuru kaynağı b ulunamadı.**'de görülebileceği gibi, genel olarak istatistiksel modelde yer alan eşitlikler ve bazı sayaçlardan oluşmaktadır.

Çizelge 5.1 Tutarlılık temelli modelden üretilen öznitelikler

Öznitelik ismi	**Açıklaması**
NEGATIVEFEEDBACK COUNTBYOVER_*t*	Bir eşin sağladığı hizmetler ile ilgili, her bir $t \in \{0.4, 0.6, 0.8\}$ eş tutarlılığı değerlerinden yüksek tutarlılığa sahip hizmet alan eşler tarafından verilen negatif değerlendirmelerin sayısını verir.
PROVIDER CONSISTENCY	Hizmet sağlayan eşin tutarlılık değeridir (Eşitlik.3)
PROVIDER FEEDBACK_[0-9]	Hizmet sağlayan eşe verilen son 10 geri bildirim
FEEDBACK_[0-9]_ COMMENTERCONS	Hizmet sağlayan eşe verilen son 10 geri bildirimi veren eşlerin tutarlılıkları
PROVIDERTRUST	Hizmet sağlayıcı eşin güven değeri (Eşitlik.8)
RECEIVER CONSTISTENCY	Hizmet alan eşin tutarlılık değeridir (Eşitlik.3)
RECEIVER CONSISTENCY_[0-9]	Hizmet alan eşin verdiği son 10 geri bildirimin tutarlılık değerleri
RECEIVERTRUST	Hizmet alan eşin güven değeri (Eşitlik.8)
TOTAL SERVICECOUNT	Hizmet sağlayıcı eşin sağlamış olduğu toplam hizmet sayısı

İthal edilen öznitelikler

Makine öğrenmesi aşamasında özniteliklerin farklı kaynaklardan elde edilmesi, farklı yaklaşımlarla elde edilmesi, makine öğrenmesi yönteminin başarım olasılığını artıracak bazı ilişkileri açığa çıkarabilir. Bu açıdan, tutarlılık temelli modelin herhangi bir noktasında karar alma sürecine katılmayan, farklı çalışmalardan esinlenilerek elde edilen ve Çizelge 5.2'de listelenen öznitelikler, bu özniteliklerin ortaya çıkarılmasını (hesaplanmasını) sağlayacak gözlemlerin benzetimden elde edilmesi ile çalışmaya dahil edilmiştir.

Çizelge 5.2 İthal edilen ya da esinlenilen öznitelikler

Öznitelik ismi	**Açıklaması (ve formülü)**
PROVIDER CHANGERATE	Gözlemlenen anlık değişimlerin oranı [50]. $$CR = \frac{\sum_{i=2}^{n} \delta_{O_{i-1},O_i}}{n-1}$$ $$\delta_{O_{i-1},O_i} = \begin{cases} 1 & if\ ((O_{i-1}=1 \wedge O_i != 1) \vee (O_{i-1} != 1 \wedge O_i = 1)) \\ 0 & Otherwise \end{cases}$$ CR hizmet veren eşin değişim oranını verir. O_i, dizideki i numaralı gözlemi temsil etmektedir.
PROVIDER TPREFIX	Başlangıç bloğunun PROVIDERTRATE [50] değeri. $$TPREFIX = \sum_{i=1}^{b} \frac{[O_i = 1]}{b}$$ Formülünde b blok büyüklüğü n/3 olarak alınmıştır. O_i, dizideki i numaralı gözlemi temsil etmektedir.
PROVIDER TRATE	Hizmet sağlayıcıya verilmiş geri bildirimler içerisindeki pozitif geri bildirim sıklığını verir [50]. $$TRATE = \sum_{i=1}^{n} \frac{[O_i = 1]}{n}$$ formülünde n geri bildirim sayısıdır. O_i, dizideki i numaralı gözlemi temsil etmektedir.
PROVIDER TRATEBYOVER_*t*	Tutarlılık değerleri $t \in \{0.4, 0.6, 0.8\}$ ‘den yüksek olan hizmet alan eşler tarafından sağlanan geri bildirimleri içeren alt kümenin Trate değeri.
PROVIDER TRIPPLE	TRIPPLE, geri bildirim dizisinde 101 örüntüsü varsa 1, yoksa 0’dır [50].
PROVIDER TSTABILITY	Dizideki değerlerin ne kadar güvenilir bir şekilde pozitif değerlerden oluştuğunu gösterir [50]. $$TSTABILITY = \frac{\sum_{i=2}^{n} \delta_{O_{i-1},O_i}}{n-1}$$ $$\delta_{O_{i-1},O_i} = \begin{cases} 1\ if\ O_{i-1} = O_i = 1 \\ 0\ Otherwise \end{cases}$$ O_i, dizideki i numaralı gözlemi temsil etmektedir.

PROVIDER TSUFFIX	Bitiş bloğunun PROVIDERTRATE değeri olarak tanımlanabilir [50]. $$TSUFFIX = \sum_{i=n-b+1}^{n} \frac{[O_i = 1]}{b}$$ O_i, dizideki i numaralı gözlemi temsil etmektedir.
PROVIDER URIPPLE	URIPPLE, geri bildirim dizisinde 010 örüntüsü varsa 1, yoksa 0'dır [50].
PROVIDER USTABILITY	Dizideki değerlerin ne kadar güvenilir bir şekilde negatif değerlerden oluştuğunu gösterir [50] $$USTABILITY = \frac{\sum_{i=2}^{n} \delta_{O_{i-1},O_i}}{n-1}$$ $$\delta_{O_{i-1},O_i} = \begin{cases} 1 \; if \; O_{i-1} = O_i = 0 \\ 0 \; Otherwise \end{cases}$$ O_i, dizideki i numaralı gözlemi temsil etmektedir.
RECEIVER CRATE	Alıcı tutarlılık oranı, 0.5 eşiği ile toplam geri bildirim sayısına karşı tutarlı geri bildirim oranını gösterir. $$R\_CRATE = \frac{\sum_{i=1}^{n} \delta_{FC_i}}{\mathrm{n}}$$ $$\delta_{FC_i} = \begin{cases} 1 \;\; FC_i > 0.5 \\ 0 \; Otherwise \end{cases}$$ formülünde FC_i *i*'nci sıradaki geri bildirimin tutarlılığını ve $n = [\mathrm{F}_r]$ değerlerini temsil etmektedir.

İkinci faz çalışmalarında üretilen öznitelikler

Makine öğrenmesinde entropi, karar ağaçlarının oluşmasında kullanılan, bilginin kazanımını (*information gain*) ölçmeye yarayan bir hesaplama yöntemidir. Entropinin ölçtüğü şey saflık ve öngörülemezliktir, 0-1 arasındaki değer 0'a yaklaştıkça saflık, 1'e yaklaştıkça öngörülemezlik artar. Çizelge 5.3'te verilen öznitelikler, Çizelge 5.1 ve Çizelge 5.2'de listelenmiş özniteliklerin ve birbirleri ile ilişkilerinin saflığını entropi ile hesaplayarak yeniden makine öğrenmesi yöntemlerine girdi olarak hazırladığımız özniteliklerdir.

Çizelge 5.3 İkinci faz çalışmaları kapsamında üretilen öznitelikler

Öznitelik ismi	**Açıklaması (ve formülü)**
CHANGEENTROPY	$(2 \times \sigma \frac{TRATE}{-CR * \log_2 CR}) - 1$ σ *Sigmoid* fonksiyonudur. 2 ile çarpıp 1 çıkarma işlemi ise özniteliğin 0.5-1 aralığından 0-1 aralığına düzeltilmesini sağlar.
ENTROPY	*Trust Rate* ile *Untrust Rate* arasındaki Entropi değeridir. $E = -(TRATE * \log_2 TRATE) - (URATE * \log_2 URATE)$ $ENTROPY = 2 \times \sigma\left(\frac{E}{URATE}\right) - 1$ σ *Sigmoid* fonksiyonudur ve $URATE = 1 - TRATE$ olarak hesaplanır. 2 ile çarpıp 1 çıkarma işlemi ise özniteliğin 0.5-1 aralığından 0-1 aralığına düzeltilmesini sağlar.
ENTROPYOF NEGATIVEFEEDBACKS	$E_{NF} = -(p1 * \log_2 p1) - (p2 * \log_2 p2)$ $ENTROPY_{NF} = 2 \times \sigma E_{NF} - 1$ *p1* değeri, Eşitlik 1 içerisindeki *satisfaction* değeri olan s'nin $s \neq 1$ olduğu durumlardaki R_{CRATE} değeridir ve $\mathrm{p2} = 1 - \mathrm{p1}$ formülü ile hesaplanır.
ENTROPYOF POSITIVEFEEDBACKS	$E_{NF} = -(p1 * \log_2 p1) - (p2 * \log_2 p2)$ $ENTROPY_{PF} = 2 \times \sigma E_{NF} - 1$ *p1* değeri, Eşitlik 1 içerisindeki *satisfaction* değeri olan s'nin $s = 1$ olduğu durumlardaki R_{CRATE} değeridir ve $\mathrm{p2} = 1 - \mathrm{p1}$ formülü ile hesaplanır.
RATEOF TRUSTCHANGERATE	Geri bildirim dizisindeki *Trust* değeri ile *Change Rate* oranını verir. $\frac{TRATE}{CR}$

5.2. Makine Öğrenimi Yöntemleri ve Öznitelik Seçimi

Daha kararlı ve güvenilir sonuçlara ulaşabilmek için 60 öznitelik ürettikten sonra, modelin başarımını artırmak için nitelik seçim algoritmaları uygulayarak seçilen öznitelikleri de deneylerde ayrıca ölçtük. Öznitelik seçimi, karmaşıklığı artıran ancak

başarıma katkı sağlamayan, çözüm için ilgisiz kalan, karmaşıklığı artırdığı için başarımı düşüren özniteliklerin azaltılmasıdır. Öznitelik seçimi, gürültülü, alakasız ve gereksiz özellikleri kaldırarak orijinal özelliklerden ilgili özelliklerin küçük bir alt kümesini seçebilir [68].

Öznitelik seçimini Weka'da yer alan *Correlation Attribute Evaluation* yöntemiyle değerlendirdik. Weka makine öğrenimi aracı dokümantasyonuna göre; "*Attribute Evaluation, bir niteliğin değerini, onunla sınıf arasındaki korelasyonu (Pearson's) ölçerek değerlendirir. Nominal nitelikler, her bir değer bir gösterge olarak ele alınarak değer bazında değerlendirilir. Nominal bir nitelik için genel bir korelasyona, ağırlıklı bir ortalama yoluyla ulaşılır* [69]. Seçilecek öznitelikleri değerlendirirken bireysel ve işbirlikçi iki yüzlü saldırgan senaryoları için benzetim çıktısı olan veri kümeleri kullanılmıştır. Bu veri kümeleri üzerinde uygulanan öznitelik seçim yöntemi çıktısı olan kümelerin birleşiminden tek bir nihai öznitelik kümesi elde edilmiştir. Bu şekilde, tekil senaryoların sınırlı doğası nedeniyle anlamlı öznitelikleri filtrelemekten kaçınmayı amaçladık. Ayrıca, hizmet ve geri bildirim saldırı senaryoları için ayrı ayrı seçim algoritmalarını çalıştırdık. Çizelge 5.4 ve Çizelge 5.5'de skorlarını verdiğimiz özniteliklerin süzülmesi için skor eşiğini deneylerimizde 0,05 ve 0,1 olarak ayrı ayrı uyguladık. 0,05 eşiğini uyguladığımız deneyler doğal olarak daha geniş bir öznitelik kümesi elde edilmektedir. Bu kümeyi uyguladığımızda elde ettiğimiz sonuçlar, 0,10 eşiğini uyguladığımızda elde ettiğimiz sonuçları iyileştiremediğinden, çizelgeleri karmaşık hale getirmemek için sonuçlara dahil etmedik ve aşağıda uygulanan deneyleri 0,10 eşiği uygulanmış öznitelik kümesi ile gerçekleştirdik.

Çizelge 5.4 Hizmet senaryoları için öznitelik skorları

PROVIDERTRATEBYOVER0_6	0,433	PROVIDERCHANGERATE	0,227
PROVIDERTRATE	0,426	PROVIDERFEEDBACK_0	0,154
PROVIDERTRATEBYOVER0_4	0,419	TOTALSERVICECOUNT	0,147
PROVIDERTRATEBYOVER0_8	0,398	NEGATIVEFEEDBACKCOUNTBYOVER0_4	0,124
PROVIDERTSUFFIX	0,397	PROVIDERFEEDBACK_1	0,120
PROVIDERTRUST	0,393	PROVIDERFEEDBACK_2	0,093
PROVIDERUSTABILITY	0,388	PROVIDERFEEDBACK_4	0,090
NEGATIVEFEEDBACKCOUNTBYOVER0_8	0,321	PROVIDERFEEDBACK_3	0,086
PROVIDERCONSISTENCY	0,293	ENTROPYOFPOSITIVEFEEDBACKS	0,078
PROVIDERTPREFIX	0,281	RECEIVERCONSTISTENCY	0,078
RATEOFTRUSTCHANGERATE	0,280	PROVIDERFEEDBACK_5	0,074
PROVIDERTRIPPLE	0,254	RECEIVERCRATE	0,074
PROVIDERTSTABILITY	0,250	PROVIDERFEEDBACK_6	0,067
CHANGEENTROPY	0,245	PROVIDERFEEDBACK_7	0,059
ENTROPY	0,232		

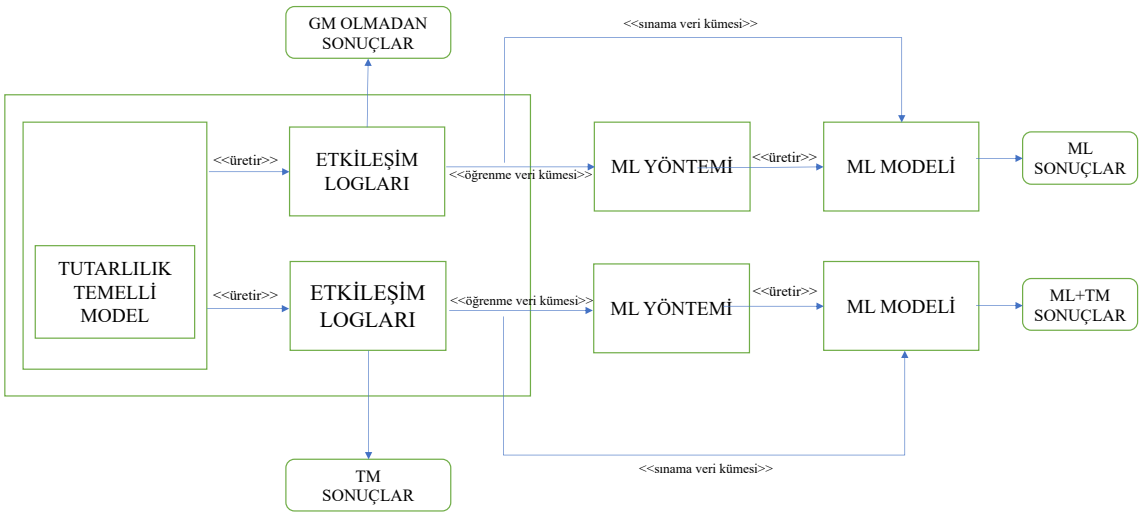
Çizelge 5.5 Geri bildirim senaryoları için öznitelik skorları

RECEIVERCRATE	0,658	PROVIDERTPREFIX	0,194
RECEIVERCONSTISTENCY	0,646	RATEOFTRUSTCHANGERATE	0,191
RECEIVERCONSISTENCY_8	0,515	CHANGEENTROPY	0,191
RECEIVERCONSISTENCY_6	0,486	PROVIDERCONSISTENCY	0,184
RECEIVERCONSISTENCY_7	0,478	PROVIDERCHANGERATE	0,179
RECEIVERCONSISTENCY_5	0,473	PROVIDERTRIPPLE	0,166
RECEIVERCONSISTENCY_3	0,468	ENTROPY	0,157
RECEIVERCONSISTENCY_2	0,459	TOTALSERVICECOUNT	0,136
RECEIVERCONSISTENCY_4	0,455	PROVIDERTSTABILITY	0,102
RECEIVERCONSISTENCY_0	0,452	PROVIDERFEEDBACK_0	0,088
RECEIVERCONSISTENCY_1	0,448	NEGATIVEFEEDBACKCOUNTBY OVER0_4	0,081
RECEIVERTRUST	0,383	PROVIDERFEEDBACK_1	0,078
PROVIDERTRATE	0,250	PROVIDERFEEDBACK_5	0,070
PROVIDERTRATEBYOVER0_6	0,246	PROVIDERFEEDBACK_3	0,066
PROVIDERUSTABILITY	0,244	PROVIDERFEEDBACK_2	0,063
PROVIDERTRATEBYOVER0_4	0,241	PROVIDERFEEDBACK_4	0,059
PROVIDERTSUFFIX	0,241	PROVIDERFEEDBACK_6	0,059
PROVIDERTRUST	0,241	PROVIDERFEEDBACK_7	0,057
PROVIDERTRATEBYOVER0_8	0,234	PROVIDERFEEDBACK_8	0,051
NEGATIVEFEEDBACKCOUNTBY OVER0_8	0,213		

5.3. Makine Öğrenmesi Modeli İçin Deneysel Çalışmalar

Peersim üzerinde geliştirdiğimiz benzetim ortamı, ağda 1000 eş aynı anda etkin olarak yapılandırılmış, toplam döngü sayısı 1000 olarak uygulanmıştır. Her senaryo için farklılık göstermekle birlikte, senaryo başına 100.000 ile 200.000 arasında değişen sayıda etkileşim verisi üretilmiştir. İstatistiksel olarak anlamlı görülen veriler benzetimin üç anında toplanmaktadır: i) etkileşim tamamlandığında, ii) döngü tamamlandığında, iii)

deney tamamlandığında. Bu istatistikler genel bir istatistik olabileceği gibi etkileşimden elde edilen bir öznitelik bileşeni üzerinden hesaplanarak elde etmek de mümkün olabilir. Bu çalışma alanındaki gerçek veri kümelerinin eksikliği sebebiyle çalışmamızda kullandığımız veri kümelerini kendi benzetim ortamımızda oluşturduk. Veri kümesinin oluşturulduğu benzetim ortamında, bu alanda yapılacak gelecek çalışmaların da bu veri oluşturma sürecinden faydalanabilmesi için makine öğrenmesi uygulamaları benzetimin içine yerleştirilmemiş, benzetimin çıktısı verilere uygulanmıştır. Benzetim verisinin oluşturulması Şekil 5.1'de gösterilmektedir.



Şekil 5.1 Benzetim verisi üretim modeli (GM: Güven Modeli, TM: Tutarlılık Modeli; ML: Makine Öğrenmesi)

Şekil 5.1 Benzetim verisi üretim modeli (GM: Güven Modeli, TM: Tutarlılık Modeli; ML: Makine Öğrenmesi) gördüğümüz akışın ilk adımında güven modelinin olmadığı ve tutarlılık temelli modelin uygulandığı deneylere ait etkileşim logları benzetim çıktısı ham veri olarak elde edilmektedir. Bu veri üzerinden güven modelinin işletilmediği bir deney ortamına ve tutarlılık temelli modelimize ait sonuçlar kayıt altına alınmaktadır. Her bir deney ikişer kez koşulmaktadır. Bu sayede hem sonuçların ortalaması alınarak daha güvenilir deney sonuçlarına ulaşılması hem de makine öğrenmesi uygulamalarında öğrenme ve sınama aşamalarında farklı veri kümelerinin uygulanması sağlanmaktadır. Makine öğrenmesi modeli hem güven modeli olmayan benzetim çıktılarına hem de tutarlılık temelli modelin benzetim çıktılarına ayrı ayrı uygulanarak sonuçlar kayıt altına alınmaktadır. Bu akış sonucunda ham etkileşim verileri, güven modelinin olmadığı deney ortamına ait sonuçlar, tutarlılık modelinin deneylerine ait sonuçlar, makine öğrenmesinin

tek başına uygulanmasından elde edilen sonuçlar ve son olarak tutarlılık temelli modelin ön işleminden geçen makine öğrenmesi uygulamalarına ait deneysel sonuçlar elde edilmektedir.

5.3.1. Saldırı Senaryoları

Tutarlılık temelli modelin sınandığı saldırgan davranışları *4.6. Saldırı Modelleri* başlığında açıklanmıştır. Makine öğrenmesi deneylerinde uygulanan bir diğer saldırgan modeli aşağıda açıklanmaktadır:

Uyarlanabilir (Adaptive) Saldırganlar: Güven değerlerini ve tutarlılık değerlerini belli bir eşiğin üstüne çıkarana kadar iyi niyetli eş olarak davranarak ağda güven sağlayan, yeterince güven sağladıktan sonra belli bir güven eşiğinin altına düşene kadar saldırarak elde ettiği güveni ağı zehirlemek için kullanan saldırgan türleridir. Örneğin güven değeri 0.8'in altında olduğu sürece iyi eş olarak davranırken, 0.8'in üstüne çıktıktan sonra durumunu değiştiren, 0.5'in altına düşene kadar saldırılar gerçekleştirmeye çalışabilirler. Önce güven verdikleri için tespit edilmeleri oldukça güçtür.

Bu tez kapsamında yalnızca hizmet sağlarken yapılan kötücül davranışlar değil, aynı zamanda geri bildirim sırasında sistemin işleyişine zarar verebilecek kötücül davranışlar da ele alınmaya çalışılmıştır. Etkileşimlerde görev alan arşivci eşlerin bu rolü yerine getirirken kötücül davranışlar göstermesi senaryoları kapsam dışında tutulmuştur. Arşivci saldırılarını değerlendirmesek de her bir eş için rastgele 3 arşivci eş tanımlayarak en az 2 arşivci eşin aynı raporu sunması güvence altına alınarak basit ancak etkili bir öneri sunulmuştur.

Bir kötücül eş saldırırken iş birliği stratejisi ile saldırı stratejisi kombinasyonu ile saldırır. Örneğin bir kötücül eş işbirlikçi toy saldırgan olabilir ya da iki yüzlü bir saldırgan olarak tek başına kötücül davranışlar gösterebilir. Bu çalışmada deneyimlenen senaryolar Çizelge 5.6'de verilmiştir.

Çizelge 5.6 Saldırı senaryoları

Kötücül Davranış	**Kötücül eş oranı**	**İş birliği stratejisi**	**Hizmet / Geribildirim**
İki yüzlü saldırgan (0.5 saldırı olasılığı ile)	%40	İşbirlikçi veya Bireysel	Hizmet veya Geri bildirim
Karma saldırgan • *İki yüzlü saldırgan (0.5 saldırı olasılığı ile)* • *Uyarlanabilir saldırgan (0.8 ile 0.2 arasında)* • *Toy saldırgan*	%15 – İki yüzlü %15 – Uyarlanabilir %15 – Toy	Bireysel	Hizmet veya Geri bildirim
Toy saldırgan	%40	İşbirlikçi veya Bireysel	Hizmet veya Geri bildirim

Makine öğrenmesi destekli modelin sınandığı deneylerde toy saldırgan oranının %20 olduğu bir senaryo da çalıştırdık, ancak sonuçları sadeleştirmek için listelediğimiz deneylerimize dahil etmedik. Bunun yerine zorlu senaryolara yer vermeyi tercih ettik. Söz konusu senaryoda, Naïve Bayes saldırıları neredeyse tamamen engelleyerek en başarılı sonuçları elde etti.

5.3.2. Makine Öğrenmesi Yapılandırmaları

Yapılan deneylerin yeniden yapılabilir olması için uygulanan makine öğrenmesi yöntemlerine ait yapılandırmalar bu başlıkta verilmektedir. Söz konusu yapılandırmalar WEKA'da uygulanabilirken, yapılandırmaya konu edilmemiş başka seçimleri WEKA'nın güncel belgelerinden bulmak mümkündür.

C45: C45 yöntemi WEKA üzerinde J48 adı ile uygulanmış bir sürümü sunar. Deneyler kapsamında uygulanan yapılandırmalar:

- *Binary splits (-B): true*

- *Pruning Confidence (-C): 0.25*
- *Number of folds (-N): 3*

SVM: SVM yöntemi WEKA üzerinde LibSVM[1] kütüphanesi ile gerçekleştirilmiştir.

- *SVM Type (-S): C-SVC*
- *Kernel function (-K): Radial basis function*
- *Degree of Kernel function (-D): 3*
- *Cache memory (-C): 40 (mb)*
- *Coef0 for kernel function (-R): 0*

Random Forest: Rasgele ağaçlar oluşturulan *Random Forest* yönteminin WEKA kütüphanesinde uygulandığı parametreler şu şekilde kullanılmıştır:

- *Number of iteration (-I): 100*
- *Minimum variance for split (-V): 0.001*
- *Number of decimal places (-num-decimal-places): 2*
- *Batch size for batch prediction (-batch-size): 100*

Isolation Forest: Aykırı durumların tespitinde kullanılmak üzere tasarlanan Isolation Forest yöntemi aşağıdaki yapılandırmalar ile kullanılmıştır:

- *Number of trees (-I): 100*
- *Size of subsample for each tree (-N): 256*
- *Number of decimal places (-num-decimal-places): 2*

Multilayer Perceptron: Deneylerde üç katmanlı olarak uygulanan Multilayer Perceptron yönteminin diğer parametreleri şu şekilde ayarlanmıştır:

[1] https://weka.sourceforge.io/doc.stable/weka/classifiers/functions/LibSVM.html

- *Learning rate (-L): 0.1*
- *Momentum (-M): 0.2*
- *Number of epoch (-N): 500*
- *Nodes for each layer (-H): 32,16*

Naïve Bayes: Bayes teoremi üzerine inşa edilen yöntemin WEKA üzerindeki gerçekleştiriminde herhangi bir parametre yapılandırılmamıştır.

5.4. Deney Yapılandırmaları

Hata! Başvuru kaynağı bulunamadı.'de verilen senaryoları farklı sınıflandırma a lgoritmaları ve öznitelik kümeleri ile irdeledik. C4.5, SVM, Random Forest, Isolation Forest, Multilayer Perceptron ve Naïve Bayes yöntemlerini yukarıda verilen yapılandırmaları ile sınadık. Öznitelikleri kaynağına veya öznitelik seçimi algoritmalarının skorlarına göre gruplandırarak deneyimledik. Kaynaklar, birinci aşamada uyguladığımız tutarlılık temelli modelimizin çıktıları/ara çıktıları, diğer çalışmalar ve makine öğrenimi modelini kullandığımız ikinci aşama oldu. Öznitelik seçiminde uyguladığımız skor eşiği sırası ile 0.1 ve 0.05 oldu. Toplamda 8 saldırı senaryosunu, 6 makine öğrenimi algoritmasını, 5 öznitelik kümesinin tüm senaryo kombinasyonlarını karşılaştırmalı olarak kaydettik. Ayrıca Şekil 5.1'de verdiğimiz 4 farklı modeli karşılaştırmaya çalıştık: herhangi bir güven modeli olmaması durumu (1), yalnızca tutarlılık temelli model (2), yalnızca makine öğrenimi modeli (3) ve birleşik model (4).

5.4.1. Değerlendirme Ölçütleri

Makine öğrenmesi yöntemlerinin ürettiği sonuçları değerlendirmek için aşağıdaki ölçütler kullanılmıştır.

Doğruluk (Accuracy): Doğru tahminlemenin toplam tahminlemeye oranıdır. Makine öğrenimi çalışmalarında kullanılan en yaygın ölçüttür

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

Hata Oranı (Error Rate): Hatalı tahminlemenin toplam tahminlemeye oranıdır.

$$Error\ Rate = 1 - Accuracy \tag{13}$$

Saldırı Oranı (Attack Rate): Saldırı yalnızca modelin etkileşime izin verdiği durumlarda yaşanır. Saldırı oranı, yanlış negatif (*false negative*) tahminlemenin toplam negatif tahminlemeye oranı ile elde edilir. Yani, saldırı olmadığı etiketlenmiş bir etkileşimin aslında kötücül bir etkileşim olması durumunun ölçümüdür.

$$Attack\ Rate = \frac{FN}{FN+TN} \tag{14}$$

5.5. Hizmet Saldırıları

5.5.1. Bireysel Saldırganlar

Toy Saldırganlar (%40)

Toy saldırganlar sürekli saldırma örüntüsü sergiledikleri için en basit, en kestirilebilir saldırganlardır. Çizelge 5.7'de görüleceği gibi, bireysel toy saldırgan senaryosunda Naïve Bayes yöntemi tutarlılık temelli model ile birlikte kullanıldığında saldırıları neredeyse tamamen engelleyen bir model olarak karşımıza çıktı. Isolation Forest dışında tüm yöntemlerin iyi bir başarım sergilediği bu saldırı senaryosunda en iyi saldırı engelleme başarımını tutarlılık temelli modelin açık olduğu senaryo için Naïve Bayes göstermiş olsa da tutarlılık temelli modelin kapalı olması durumunda Random Forest en iyi, MLP ise ikinci en iyi başarımı göstermektedir (Çizelge 5.8).

Çizelge 5.7 Tutarlılık temelli filtreleme açık olduğunda bireysel toy saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	99,56%	0,44%	0,34%	99,56%	0,44%	0,35%	99,57%	0,43%	0,35%	99,57%	0,43%	0,35%
SVM	99,40%	0,60%	0,45%	99,42%	0,58%	0,44%	99,48%	0,52%	0,39%	99,42%	0,58%	0,45%
Random Forest	99,56%	0,44%	0,34%	99,56%	0,44%	0,33%	99,56%	0,44%	0,33%	99,55%	0,45%	0,35%
MLP	99,48%	0,52%	0,34%	99,51%	0,49%	0,37%	99,51%	0,49%	0,37%	99,50%	0,50%	0,40%
Naive Bayes	95,65%	4,35%	0,00%	98,54%	1,46%	0,00%	99,17%	0,83%	0,00%	99,17%	0,83%	0,00%
Isolation Forest	89,78%	10,22%	0,22%	88,82%	11,18%	0,08%	88,42%	11,58%	0,00%	86,26%	13,74%	0,00%

Çizelge 5.8 Tutarlılık temelli filtreleme kapalı olduğunda bireysel toy saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	99,50%	0,50%	1,03%	99,53%	0,47%	1,08%	99,53%	0,47%	1,08%	99,51%	0,49%	0,92%
SVM	98,97%	1,03%	1,55%	99,52%	0,48%	0,88%	99,52%	0,48%	0,88%	99,52%	0,48%	0,88%
Random Forest	99,54%	0,46%	0,84%	99,53%	0,47%	0,86%	99,53%	0,47%	0,86%	99,49%	0,51%	0,89%
MLP	99,52%	0,48%	0,97%	99,52%	0,48%	0,95%	99,54%	0,46%	0,95%	99,52%	0,48%	0,85%
Naive Bayes	97,27%	2,73%	7,52%	99,20%	0,80%	1,97%	99,40%	0,60%	1,32%	99,51%	0,49%	0,90%
Isolation Forest	25,93%	74,07%	75,44%	25,80%	74,20%	75,56%	23,50%	76,50%	77,44%	29,88%	70,12%	75,15%

Tutarlılık temelli modelin açık ya da kapalı olduğu her iki senaryo için de Naïve Bayes yönteminin, Isolation Forest dışındaki tüm algoritmalardan genelde daha fazla hata oranı ürettiğini gözlemledik. Bu durumu Naïve Bayes modelinin, saldırganların bir kısmını normal eş olarak sınıflamaya daha meyilli olmasından kaynaklanmaktadır. Fakat hata oranı çok yüksek olmadığı için, bu senaryoda Naïve Bayes modeli genel olarak başarılı bir model olarak öne çıkmaktadır.

İki yüzlü saldırgan senaryosu (0.5 saldırı olasılığı ile)

İki yüzlü saldırgan senaryosunu, %40 saldırgan ve 0,5 saldırı oranı ile uygulandığımızda, herhangi güven modeli kullanılmayınca %20 oranında bir saldırı beklenir. Tutarlılık temelli istatistiksel yaklaşımımız, bu senaryoda ağdaki saldırı oranını tek başına %1,29 oranına kadar indirebilmektedir. Bu yüksek başarımı makine öğrenimi algoritmaları ile desteklediğimizde, Çizelge 5.9'da görülebileceği gibi, saldırı oranının %0,01 oranlarına kadar iyileştirildiğini gözlemledik. Bu da bize yöntemimizin makine öğrenimi ile birlikte

bu senaryo için ne kadar uyumlu olduğunu gösterdi. Sonuçların ayrıntısına bakarsak, Naïve Bayes ve Isolation Forest yöntemlerinin saldırıları engellemede benzer iyi başarımları sergilediğini, Naïve Bayes'in doğruluk ölçütünde Isolation Forest'a göre daha iyi olduğunu gözlemledik. Ne var ki, diğer yöntemlerle karşılaştırdığımızda Naïve Bayes yöntemi, bütün etkileşimleri saldırı olarak etiketlemeye daha duyarlı olduğu için saldırı engellemekte başarılıdır. Fakat model, normal etkileşimleri de saldırı olarak sınıflandırma eğiliminden ötürü genel doğruluk ölçütünde daha düşük başarım göstermektedir. Eğer kurulmak istenen eşler arası ağın gereksinimi doğruluğa saldırı oranından daha çok değer veriyor ise Random Forest ya da MLP yöntemleri ile eğitilen bir modelin kullanılması daha iyi bir seçim olabilir.

Çizelge 5.9 Tutarlılık temelli filtreleme açık olduğunda bireysel iki yüzlü saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	98,83%	1,17%	0,89%	98,90%	1,10%	0,81%	98,90%	1,10%	0,80%	98,83%	1,17%	0,81%
SVM	98,85%	1,15%	1,15%	94,84%	5,16%	1,91%	98,85%	1,15%	1,15%	98,85%	1,15%	1,15%
Random Forest	98,88%	1,12%	0,87%	98,96%	1,04%	0,80%	98,94%	1,06%	0,81%	98,81%	1,19%	0,79%
MLP	98,87%	1,13%	0,86%	98,86%	1,14%	0,78%	98,88%	1,12%	0,80%	98,85%	1,15%	0,94%
Naive Bayes	94,74%	5,26%	0,06%	95,15%	4,85%	0,02%	94,96%	5,04%	0,01%	95,32%	4,68%	0,01%
Isolation Forest	83,23%	16,77%	0,02%	84,71%	15,29%	0,00%	84,52%	15,48%	0,00%	88,18%	11,82%	0,00%

Bireysel iki yüzlü saldırgan senaryosu için tutarlılık temelli filtrelemeyi kapattığımız durumda Naive Bayes'in filtrenin açık olduğu senaryoda olduğu gibi diğer yöntemlerden daha yüksek bir başarım ile saldırıları engellediğini gözlemledik (Çizelge 5.10). iki yüzlü saldırgan senaryolarında tutarlılık modelinin açık olduğu makine öğrenimi modellerinin genel olarak daha başarılı sonuçlar elde etmemizi sağladığını söyleyebiliriz. Ayrıca tutarlılık temelli modelinin açık olduğu senaryolarda Naïve Bayes saldırı engellemede diğer yöntemlerden çok daha iyi sonuçlar elde ederken, tutarlılık temelli modeli kapattığımızda saldırı engellemede yine diğer yöntemlerden daha başarılı olduğunu gözlemledik. Diğer senaryolardaki genel izlenimden farklı olarak, Naive Bayes algoritmasının bu kez doğruluk ölçütünde diğer senaryolara yakın sonuç ürettiğini gördük.

Çizelge 5.10 Tutarlılık temelli filtreleme kapalı olduğunda bireysel iki yüzlü saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	84,03%	15,97%	8,60%	84,58%	15,42%	9,03%	84,63%	15,37%	9,04%	84,00%	16,00%	1,76%
SVM	83,83%	16,17%	12,20%	91,25%	8,75%	8,75%	84,48%	15,52%	6,51%	84,00%	16,00%	5,59%
Random Forest	84,29%	15,71%	9,91%	84,85%	15,15%	9,48%	84,82%	15,18%	9,46%	84,53%	15,47%	9,44%
MLP	83,89%	16,11%	8,31%	84,34%	15,66%	7,76%	84,39%	15,61%	7,67%	83,76%	16,24%	9,36%
Naive Bayes	83,72%	16,28%	6,08%	83,82%	16,18%	1,39%	83,81%	16,19%	1,52%	83,92%	16,08%	1,45%
Isolation Forest	75,87%	24,13%	9,58%	76,37%	23,63%	10,28%	76,27%	23,73%	10,89%	75,31%	24,69%	10,76%

Karma Saldırganlar

Karma saldırı senaryolarında, Naïve Bayes saldırı oranı açısından en güvenilir modeli sağlarken en yüksek hata oranına ulaşmaktadır (Çizelge 5.11). Sonuçlardan anlaşılacağı üzere Isolation Forest yönteminin bizim benzetim ortamımız için çok düzensiz, güvenilmez sonuçlar verdiği görülmektedir. Tutarlılık temelli modelin kapatılarak yalnızca makine öğrenimi modelleri ile deneyleri koştuğumuzda Naïve Bayes yine kötücül davranışlara karşı en büyük korumayı sağlarken Random Forest de iyi başarım gösteren bir diğer yöntem oldu (Çizelge 5.12).

Çizelge 5.11 Tutarlılık temelli filtreleme açık olduğunda bireysel karma (%15 iki yüzlü, %15 toy, %15 uyarlanabilir) saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	99,01%	0,99%	0,74%	99,50%	0,50%	0,74%	98,98%	1,02%	0,72%	99,03%	0,97%	0,73%
SVM	98,73%	1,27%	1,27%	93,64%	6,36%	5,92%	98,98%	1,02%	0,97%	98,97%	1,03%	0,99%
Random Forest	99,03%	0,97%	0,83%	99,67%	0,33%	0,16%	99,06%	0,94%	0,79%	99,05%	0,95%	0,75%
MLP	98,95%	1,05%	0,83%	99,00%	1,00%	0,69%	98,98%	1,02%	0,71%	99,03%	0,97%	0,78%
Naive Bayes	93,34%	6,66%	0,11%	94,04%	5,96%	0,08%	93,64%	6,36%	0,07%	94,31%	5,69%	0,07%
Isolation Forest	13,45%	86,55%	8,40%	12,10%	87,90%	9,43%	12,58%	87,42%	9,13%	13,38%	86,62%	8,64%

Çizelge 5.12 Tutarlılık temelli filtreleme kapalı olduğunda bireysel karma (%15 iki yüzlü, %15 toy, %15 uyarlanabilir) saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	93,75%	6,25%	5,50%	94,07%	5,93%	5,99%	94,22%	5,78%	6,22%	94,28%	5,72%	6,54%
SVM	93,49%	6,51%	6,12%	83,45%	16,55%	15,85%	94,36%	5,64%	6,59%	94,23%	5,77%	6,71%
Random Forest	93,80%	6,20%	5,61%	93,99%	6,01%	5,73%	94,00%	6,00%	5,70%	93,91%	6,09%	5,08%
MLP	93,98%	6,02%	5,97%	94,08%	5,92%	6,07%	94,07%	5,93%	6,03%	94,32%	5,68%	6,64%
Naive Bayes	92,16%	7,84%	3,05%	90,89%	9,11%	3,48%	91,02%	8,98%	3,61%	90,57%	9,43%	3,66%
Isolation Forest	72,53%	27,47%	17,87%	71,35%	28,65%	17,05%	72,20%	27,80%	16,43%	74,50%	25,50%	8,17%

5.5.2. İşbirlikçi Saldırganlar

Toy Saldırganlar (%40)

Toy saldırganların iş birliği yapması durumunda, Isolation Forest tutarlılık temelli model ile ağdaki saldırıları en iyi şekilde engellerken, doğruluk ölçütüne göre ise en kötü başarımı gösterdi (Çizelge 5.13). C45, SVM, Random Forest ve MLP kendi aralarında benzer doğrulukları gösterirken MLP saldırıları önlemede bu grup içinde en iyisidir. Naïve Bayes, bu senaryoda da güvenilir bir önleme sağlıyor. Random Forest yöntemini, ithal edilen öznitelik kümesiyle ve tüm özniteliklerle kullandığımızda veya MLP yöntemini tüm özniteliklerle kullandığımızda elde ettiğimiz çıktılar, genel görünümde en iyi sonuçları gösteriyor. Sonuçlardan anlaşılacağı üzere, Isolation Forest dışındaki tüm yöntemler benzer düzeyde iyi sonuçları elde etmeyi başarıyor.

Tutarlılık temelli modeli devre dışı bıraktığımızda Isolation Forest yöntemi hem saldırı engellemede hem de doğruluk ölçütünde diğer yöntemlerin çok daha gerisinde başarımlar sergilemektedir. Diğer yöntemlerden MLP'nin tek başına özellikle tüm öznitelikler ile birlikte uygulandığında bir miktar daha iyi sonuçlar ürettiğini, Random Forest'ın ise ithal öznitelikler ile uygulandığında en doğru sınıflandırma başarımına ulaştığını gözlemledik.

Çizelge 5.13 Tutarlılık temelli filtreleme açık olduğunda işbirlikçi toy saldırgan senaryo sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	98,59%	1,41%	1,02%	98,60%	1,40%	1,01%	98,64%	1,36%	1,01%	98,59%	1,41%	0,99%
SVM	98,35%	1,65%	1,17%	98,46%	1,54%	1,10%	98,50%	1,50%	1,04%	98,50%	1,50%	1,04%
Random Forest	98,72%	1,28%	0,99%	98,75%	1,25%	0,95%	98,74%	1,26%	0,94%	98,55%	1,45%	0,99%
MLP	98,50%	1,50%	0,88%	98,62%	1,38%	0,96%	98,61%	1,39%	0,90%	98,59%	1,41%	0,98%
Naive Bayes	92,18%	7,82%	0,38%	92,18%	7,82%	0,03%	93,66%	6,34%	0,09%	95,38%	4,62%	0,17%
Isolation Forest	88,24%	11,76%	0,04%	87,66%	12,34%	0,00%	87,52%	12,48%	0,00%	86,38%	13,62%	0,00%

Çizelge 5.14 Tutarlılık temelli filtreleme kapalı olduğunda işbirlikçi toy saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	98,71%	1,29%	1,36%	98,89%	1,11%	1,30%	98,88%	1,12%	1,28%	98,71%	1,29%	1,24%
SVM	98,31%	1,69%	1,86%	98,75%	1,25%	1,29%	98,78%	1,22%	1,31%	98,52%	1,48%	1,29%
Random Forest	98,94%	1,06%	1,25%	99,01%	0,99%	1,15%	98,98%	1,02%	1,15%	98,68%	1,32%	1,29%
MLP	98,68%	1,32%	1,34%	98,89%	1,11%	1,28%	98,93%	1,07%	1,12%	98,70%	1,30%	1,16%
Naive Bayes	97,85%	2,15%	1,86%	98,29%	1,71%	1,20%	98,25%	1,75%	1,76%	98,25%	1,75%	1,75%
Isolation Forest	60,95%	39,05%	38,77%	58,97%	41,03%	40,07%	58,32%	41,68%	40,17%	58,49%	41,51%	38,55%

İki yüzlü saldırgan senaryosu (0.5 saldırı oranı ile)

Bireysel saldırgan modelindeki iki yüzlü saldırgan senaryosuna benzer şekilde, iş birliği yapan iki yüzlü kötücül eşler için de deneyler yaptık. Bu deneyler, modelimizin bireysel olanlar kadar işbirlikçi kötü niyetli eşleri tespit etmede başarılı olduğunu göstermektedir (Çizelge 5.15). Naïve Bayes, seçilmiş öznitelikler ya da tüm özniteliklerle saldırılara karşı en iyi önleyici yöntem olmuştur. Bununla birlikte, doğruluk daha önemli bir ölçüt olarak belirlenirse, Naïve Bayes bu senaryo için de en iyi sonucu veriyor denemez. Bu tür bir öncelik geçerli ise senaryoda MLP algoritmasının daha doğru sonuçlar verdiğini söyleyebiliriz.

Çizelge 5.15 Tutarlılık temelli filtreleme açık olduğunda işbirlikçi iki yüzlü saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	97,05%	2,95%	1,99%	97,04%	2,96%	1,90%	97,00%	3,00%	2,00%	97,13%	2,87%	2,87%
SVM	97,13%	2,87%	2,87%	97,13%	2,87%	2,87%	97,13%	2,87%	2,87%	97,13%	2,87%	2,87%
Random Forest	97,19%	2,81%	2,23%	97,24%	2,76%	2,17%	97,20%	2,80%	2,19%	96,63%	3,37%	2,69%
MLP	97,17%	2,83%	1,85%	97,15%	2,85%	1,80%	97,15%	2,85%	1,67%	97,13%	2,87%	2,87%
Naive Bayes	90,78%	9,22%	0,75%	90,37%	9,63%	0,57%	90,39%	9,61%	0,49%	90,30%	9,70%	0,50%
Isolation Forest	81,82%	18,18%	0,89%	84,06%	15,94%	0,66%	85,91%	14,09%	0,43%	87,38%	12,62%	0,29%

Tutarlılık temelli filtreleme adımı kapatıldığında, Naïve Bayes, tüm yöntemler arasında sistemi saldırılara karşı korumada yine en iyi performansı gösteriyor (Çizelge 5.16). Buna rağmen sonuçlar, ML algoritmalarını çalıştırmadan önce tutarlılık temelli modeli etkinleştirmenin daha iyi olduğunu göstermektedir.

Çizelge 5.16 Tutarlılık temelli filtreleme kapalı olduğunda işbirlikçi iki yüzlü saldırgan senaryosu için hizmet saldırıları sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	87,29%	12,71%	7,35%	87,83%	12,17%	7,05%	87,85%	12,15%	47,91%	87,25%	12,75%	12,65%
SVM	87,31%	12,69%	12,69%	87,31%	12,69%	12,69%	87,31%	12,69%	12,69%	87,31%	12,69%	12,69%
Random Forest	87,71%	12,29%	8,06%	88,20%	11,80%	7,54%	88,17%	11,83%	7,69%	85,20%	14,80%	11,63%
MLP	87,47%	12,53%	7,11%	87,87%	12,13%	7,04%	87,83%	12,17%	6,84%	87,31%	12,69%	12,69%
Naive Bayes	85,43%	14,57%	5,54%	83,59%	16,41%	3,16%	78,73%	21,27%	2,52%	76,46%	23,54%	2,64%
Isolation Forest	27,40%	72,60%	18,30%	26,65%	73,35%	17,82%	24,54%	75,46%	16,63%	22,52%	77,48%	21,88%

5.6. Geri Bildirim Saldırıları

5.6.1. Bireysel Saldırganlar

Toy Saldırganlar (%40)

Toy bireysel saldırganlar için çizelgelerin genel görünümünü gösteren Çizelge 5.17 ve Çizelge 5.18'ü karşılaştırdığımızda; makine öğrenimi yöntemleri, toy bireysel geri bildirim saldırganlarına karşı tutarlılık temelli modelle birlikte kullanıldığında sistemleri daha iyi korumaktadır. Öznitelik kümesi olarak, ithal edilen öznitelikleri veya tüm öznitelikleri seçtiğimizde, algoritmaların doğruluk ve saldırı oranı ölçütlerinde daha iyi

olduğu gözlemlenmiştir. Tutarlılık temelli model kullanıldığında, Naive Bayes modeli bu saldırgan modeli için geri bildirim saldırılarını önlemede hizmet saldırılarında olduğu gibi en başarılı model olmaktadır.

Çizelge 5.17 Tutarlılık temelli filtreleme açık olduğunda bireysel toy saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	98,86%	1,14%	0,71%	99,02%	0,98%	0,69%	99,06%	0,94%	0,67%	97,70%	2,30%	2,14%
SVM	98,61%	1,39%	1,01%	98,70%	1,30%	0,92%	98,73%	1,27%	0,89%	98,75%	1,25%	0,92%
Random Forest	99,05%	0,95%	0,69%	99,15%	0,85%	0,63%	99,16%	0,84%	0,62%	97,51%	2,49%	2,02%
MLP	98,71%	1,29%	0,73%	98,90%	1,10%	0,60%	98,92%	1,08%	0,64%	97,70%	2,30%	2,09%
Naive Bayes	91,25%	8,75%	0,13%	91,17%	8,83%	0,01%	92,04%	7,96%	0,02%	93,16%	6,84%	0,55%
Isolation Forest	88,89%	11,11%	0,05%	88,56%	11,44%	0,02%	88,30%	11,70%	0,00%	86,35%	13,65%	0,15%

Çizelge 5.18 Tutarlılık temelli filtreleme kapalı olduğunda bireysel toy saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	80,90%	19,10%	12,50%	99,39%	0,61%	0,67%	99,39%	0,61%	0,68%	63,00%	37,00%	36,55%
SVM	75,77%	24,23%	17,00%	99,10%	0,90%	0,94%	99,11%	0,89%	0,94%	98,89%	1,11%	1,61%
Random Forest	75,39%	24,61%	19,39%	99,28%	0,72%	0,78%	99,28%	0,72%	0,78%	57,99%	42,01%	35,64%
MLP	76,46%	23,54%	14,03%	99,19%	0,81%	0,94%	99,22%	0,78%	0,89%	62,68%	37,32%	36,23%
Naive Bayes	73,05%	26,95%	17,33%	99,05%	0,95%	1,03%	98,98%	1,02%	1,27%	49,28%	50,72%	27,83%
Isolation Forest	42,66%	57,34%	35,65%	53,25%	46,75%	29,08%	55,67%	44,33%	37,08%	55,99%	44,01%	38,65%

İki yüzlü saldırgan senaryosu (0.5 saldırı oranı ile)

İki yüzlü saldırgan senaryosunu, %40 saldırgan ve 0,5 saldırı oranı ile uygulandığımızda, herhangi güven modeli kullanılmayınca %20 oranında bir saldırı beklenir. Tutarlılık temelli istatistiksel yaklaşımımız, bu senaryoda ağdaki saldırı oranını tek başına %6,89 oranına kadar indirebilmektedir. Çizelge 5.19 ve Çizelge 5.20 raporladığımız deneysel sonuçlar bize, modelimizi makine öğrenmesi yöntemleri ile birleştirmenin genel görünümü iyileştirdiğini gösteriyor.

Çizelge 5.19 Tutarlılık temelli filtreleme açık olduğunda bireysel iki yüzlü saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	94,94%	5,06%	2,83%	94,95%	5,05%	2,83%	94,95%	5,05%	2,80%	98,96%	1,04%	0,72%
SVM	94,82%	5,18%	2,15%	94,84%	5,16%	1,91%	94,87%	5,13%	1,91%	94,87%	5,13%	1,65%
Random Forest	95,05%	4,95%	2,94%	95,17%	4,83%	2,82%	95,11%	4,89%	2,85%	98,99%	1,01%	0,75%
MLP	94,98%	5,02%	2,58%	94,83%	5,17%	2,51%	94,80%	5,20%	2,61%	95,07%	4,93%	2,59%
Naive Bayes	92,92%	7,08%	0,56%	91,85%	8,15%	0,43%	91,65%	8,35%	0,40%	95,09%	4,91%	0,01%
Isolation Forest	85,87%	14,13%	1,52%	85,87%	14,13%	2,18%	85,30%	14,70%	2,55%	81,87%	18,13%	0,00%

Çizelge 5.20'de sonuçları verilen tutarlılık temelli modelin açık olduğu deneyler için Naïve Bayes, tüm öznitelik grupları ile saldırıları engellemede en başarılı model olarak ön plana çıkmaktadır ve seçili öznitelikler ile uygulandığında geri bildirim saldırı oranlarını %0,01 oranına kadar azaltmayı başarmaktadır. Doğruluk için ise en iyi sonuçları seçili öznitelik kümesini uyguladığımızda Random Forest ve C45 algoritmaları ile eğitilen modellerden elde ettik. Bu modellerin saldırı oranları da %1'in altında olduğu için, bu modellerin Naive Bayes'a alternatif olması mümkün görünmektedir.

Çizelge 5.20 Tutarlılık temelli filtreleme kapalı olduğunda bireysel iki yüzlü saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	79,58%	20,42%	19,00%	82,26%	17,74%	10,04%	82,28%	17,72%	10,10%	99,12%	0,88%	0,72%
SVM	79,66%	20,34%	20,34%	80,83%	19,17%	5,52%	80,79%	19,21%	5,44%	80,79%	19,21%	4,89%
Random Forest	76,55%	23,45%	17,96%	80,96%	19,04%	12,69%	80,86%	19,14%	12,77%	99,08%	0,92%	0,66%
MLP	79,66%	20,34%	20,34%	80,80%	19,20%	9,37%	80,79%	19,21%	9,36%	80,87%	19,13%	9,37%
Naive Bayes	73,40%	26,60%	13,10%	80,32%	19,68%	3,06%	80,30%	19,70%	3,15%	95,72%	4,28%	0,03%
Isolation Forest	53,97%	46,03%	21,30%	68,61%	31,39%	20,18%	69,66%	30,34%	20,24%	11,88%	88,12%	7,27%

Bireysel iki yüzlü saldırganların geri bildirim saldırıları, tutarlılık modeli filtrelemesi olmadan seçili öznitelikler kullanılınca benzeri başarım ile tespit edilebilmektedir (Çizelge 5.20). Ayrıca, tüm iki yüzlü bireysel geri bildirim saldırgan senaryolarında en iyi doğruluk değerlerini, seçili öznitelikleri uyguladığımızda C45 ve Random Forest yöntemlerinin ürettiğini gözlemledik. Öznitelik kümesinin skorlanarak seçilmesinin bu

senaryodaki deneylerde tüm yöntemlerin genel başarımlarını iyileştirdiğini söyleyebiliriz.

Karma Saldırganlar

Karma saldırı senaryosu altında, Naïve Bayes, hizmet saldırısı deneylerinde olduğu gibi en güvenilir ağı sağlayabiliyor (Çizelge 5.21). Öte yandan, Random Forest ve C45 benzer şekilde doğruluk ölçütü için daha iyi sonuçlar almamızı sağlıyor. Tüm yöntemlerin en yüksek başarımı seçili öznitelik kümesinin uygulandığı deneylerde gösterdiğini gözlemledik.

Çizelge 5.21 Tutarlılık temelli filtreleme açık olduğunda bireysel karma (%15 iki yüzlü, %15 toy, %15 uyarlanabilir) saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	93,79%	6,21%	4,43%	93,89%	6,11%	4,40%	93,87%	6,13%	4,34%	98,99%	1,01%	0,68%
SVM	93,41%	6,59%	6,39%	93,64%	6,36%	5,92%	93,62%	6,38%	5,97%	93,60%	6,40%	6,02%
Random Forest	94,04%	5,96%	4,44%	94,06%	5,94%	4,42%	94,15%	5,85%	4,42%	99,03%	0,97%	0,81%
MLP	93,94%	6,06%	4,26%	94,01%	5,99%	4,38%	93,96%	6,04%	4,26%	94,16%	5,84%	4,75%
Naive Bayes	88,93%	11,07%	1,03%	88,83%	11,17%	1,04%	88,31%	11,69%	1,21%	93,90%	6,10%	0,07%
Isolation Forest	86,90%	13,10%	3,32%	87,28%	12,72%	3,96%	85,97%	14,03%	3,93%	15,38%	84,62%	7,59%

Çizelge 5.22 Tutarlılık temelli filtreleme kapalı olduğunda bireysel karma (%15 iki yüzlü, %15 toy, %15 uyarlanabilir) saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	74,15%	25,85%	23,69%	91,15%	8,85%	10,07%	91,13%	8,87%	10,08%	99,21%	0,79%	0,72%
SVM	67,14%	32,86%	32,86%	90,87%	9,13%	10,92%	90,89%	9,11%	10,90%	90,88%	9,12%	10,92%
Random Forest	67,79%	32,21%	26,43%	90,60%	9,40%	9,66%	90,64%	9,36%	9,73%	99,15%	0,85%	0,70%
MLP	68,23%	31,77%	28,33%	90,86%	9,14%	10,65%	90,79%	9,21%	10,68%	90,95%	9,05%	10,81%
Naive Bayes	66,03%	33,97%	24,33%	89,35%	10,65%	5,55%	89,13%	10,87%	5,78%	94,61%	5,39%	0,15%
Isolation Forest	49,69%	50,31%	32,87%	61,11%	38,89%	31,31%	61,75%	38,25%	31,57%	84,34%	15,66%	0,05%

Tutarlılık temelli model etkileşimleri filtrelemediğinde makine öğrenmesi yöntemlerinin daha az başarılı olduğunu söyleyebiliriz (Çizelge 5.22). Bu ortalama bir gözlem olması

açısından güçlü bir gözlem olsa da, seçilen özniteliklerin yöntemler üzerinde olumlu etkiler yarattığını gözlemlediğimizi atlamamalıyız. Isolation Forest'ın saldırı oranı başarımını %30'lardan %0.05'e çarpıcı bir şekilde iyileştirdiğini gözlemledik. Ancak, Isolation Forest'ın doğruluk ölçütü değerinin, seçilen öznitelik kümesi ile sınanan yöntemler arasında en düşük başarımı gösterdiğini de not etmemiz gerekir. Öte yandan, C45 algoritmasının seçili öznitelik kümesi ile birlikte uygulandığında en iyi doğruluğu verdiğini gözlemledik.

5.6.2. İşbirlikçi Saldırganlar

Toy Saldırganlar (%40)

Saldırganlar sistem içinde iş birliği gösterirse, tutarlılık temelli modeli açmak ile kapatmak deneylerinin sonuçları üzerinde büyük farklılıklara sebep olabilmektedir. Deney sonuçlarından Isolation Forest, toy saldırganların iş birliğine dayalı geri bildirim saldırganlarına karşı sisteme neredeyse tam koruma sağlayabildiğini söyleyebiliriz, fakat bunu yaparken normal geribildirimleri de saldırı olarak algılamaya meyilli olduğu için doğruluk oranı diğer yöntemlerden düşüktür. Bu nedenle, düşük saldırı oranı ve yüksek doğruluk oranı sağlama açısından Naive Bayes modeli tutarlılık temelli modelle birlikte kullanıldığında en başarılı olandır. Çizelge 5.23 ve Çizelge 5.24'daki sonuçlara göre, ithal edilen öznitelik kümesi ve tüm öznitelikler kümesi, toy saldırganlar için daha uygundur. Ayrıca çizelgelerde, tutarlılık temelli filtreleme olmadan uygulanan yöntemlerin ve öznitelik kümelerinin sonuçlarda çok büyük etkileri olduğu görülebilir (Çizelge 5.24). Tutarlılık temelli model kullanılmadığında C.45 ve Random Forest algoritmaları öne çıkmaktadır.

Çizelge 5.23 Tutarlılık temelli filtreleme açık olduğunda işbirlikçi toy saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	97,73%	2,27%	1,54%	97,80%	2,20%	1,45%	97,82%	2,18%	1,48%	97,05%	2,95%	2,57%
SVM	97,10%	2,90%	2,76%	97,10%	2,90%	2,72%	97,09%	2,91%	2,70%	96,95%	3,05%	3,05%
Random Forest	97,96%	2,04%	1,57%	98,03%	1,97%	1,49%	98,02%	1,98%	1,52%	96,86%	3,14%	2,38%
MLP	97,63%	2,37%	1,50%	97,66%	2,34%	1,33%	97,65%	2,35%	1,36%	97,02%	2,98%	2,55%
Naive Bayes	90,39%	9,61%	0,46%	90,43%	9,57%	0,08%	91,65%	8,35%	0,15%	93,12%	6,88%	0,83%
Isolation Forest	88,18%	11,82%	0,06%	87,61%	12,39%	0,02%	87,26%	12,74%	0,01%	85,87%	14,13%	0,15%

Çizelge 5.24 Tutarlılık temelli filtreleme kapalı olduğunda işbirlikçi toy saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	80,28%	19,72%	16,89%	97,12%	2,88%	3,21%	97,14%	2,86%	3,14%	63,78%	36,22%	36,06%
SVM	67,47%	32,53%	28,29%	94,59%	5,41%	4,30%	94,72%	5,28%	4,34%	96,85%	3,15%	3,58%
Random Forest	72,10%	27,90%	23,27%	97,06%	2,94%	3,24%	97,03%	2,97%	3,31%	59,03%	40,97%	34,56%
MLP	70,25%	29,75%	20,17%	96,75%	3,25%	3,54%	96,70%	3,30%	3,57%	63,79%	36,21%	36,21%
Naive Bayes	67,22%	32,78%	20,25%	94,30%	5,70%	5,13%	94,31%	5,69%	5,82%	49,56%	50,44%	25,32%
Isolation Forest	40,81%	59,19%	40,50%	48,94%	51,06%	35,31%	50,32%	49,68%	38,97%	51,23%	48,77%	39,80%

İki yüzlü saldırgan senaryosu (0.5 saldırı oranı ile)

İki yüzlü işbirlikçi saldırganların geri bildirim saldırılarının tespiti, sonuç tablosunda da görülebileceği gibi daha zor hale gelmektedir (Çizelge 5.25). En iyi önlemenin gerçekleştiği gözlemlerde bile, tüm özniteliklerin kullanıldığı Naive Bayes gözleminde, saldırı oranı, Çizelge 5.19'de yer alan en başarılı saldırı oranı olan %0.01'den %1.26'ya yükselmektedir. Ayrıca, işbirlikçi saldırgan senaryolarında seçili öznitelik kümesinin uygulanması, bireysel saldırgan senaryolarının aksine en kötü performansı gösterebilmektedir.

Çizelge 5.25 Tutarlılık temelli filtreleme açık olduğunda işbirlikçi iki yüzlü saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	90,94%	9,06%	5,02%	90,90%	9,10%	4,80%	90,77%	9,23%	4,76%	89,97%	10,03%	10,03%
SVM	90,90%	9,10%	5,75%	90,93%	9,07%	5,77%	91,00%	9,00%	5,57%	89,97%	10,03%	10,03%
Random Forest	91,12%	8,88%	5,57%	91,18%	8,82%	5,44%	91,07%	8,93%	5,56%	87,27%	12,73%	10,14%
MLP	91,03%	8,97%	4,52%	91,03%	8,97%	4,64%	90,99%	9,01%	4,52%	89,97%	10,03%	10,03%
Naive Bayes	88,31%	11,69%	1,54%	88,78%	11,22%	1,32%	88,95%	11,05%	1,26%	29,44%	70,56%	9,77%
Isolation Forest	79,14%	20,86%	7,62%	79,51%	20,49%	8,25%	77,26%	22,74%	8,41%	78,67%	21,33%	10,22%

Deneylerimizden elde ettiğimiz sonuçlara göre, makine öğrenmesi yöntemleri, iki yüzlü saldırganların iş birliği yaptıkları senaryoda, tutarlılık temelli model olmadan daha kötü başarım göstermektedir (Çizelge 5.26). Bu senaryoda, tüm öznitelikler kümesini uyguladığımız Naïve Bayes'in en iyi saldırı önleme başarımını gösterdiğini gözlemledik. İlginç bir şekilde, seçilen öznitelikleri uyguladığımız Naïve Bayes'in ise oldukça kötü doğruluk sonuçları verdiğini gözlemledik. Sonuçlar bize, özniteliklerin en iyi alt kümesini seçmenin sonucu büyük ölçüde etkilediğini gösteriyor.

Çizelge 5.26 Tutarlılık temelli filtreleme kapalı olduğunda işbirlikçi iki yüzlü saldırgan senaryosu için geri bildirim saldırısı sonuçları

Yöntem	Temel Öznitelikler			İthal Öznitelikler + Temel Öznitelikler			Tüm Öznitelikler			Seçili Öznitelikler (>0.1)		
	Doğruluk	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*	*Doğruluk*	*Hata Oranı*	*Saldırı Oranı*
C45	82,08%	17,92%	16,76%	83,80%	16,20%	9,47%	83,74%	16,26%	9,49%	82,46%	17,54%	17,54%
SVM	82,46%	17,54%	17,54%	82,58%	17,42%	17,15%	82,66%	17,34%	16,43%	82,46%	17,54%	17,54%
Random Forest	79,68%	20,32%	17,05%	83,35%	16,65%	11,53%	83,13%	16,87%	11,85%	77,53%	22,47%	17,64%
MLP	82,46%	17,54%	17,54%	83,26%	16,74%	9,62%	83,26%	16,74%	9,32%	82,46%	17,54%	17,54%
Naive Bayes	82,46%	17,54%	17,54%	82,89%	17,11%	4,46%	82,71%	17,29%	3,69%	41,02%	58,98%	13,46%
Isolation Forest	52,27%	47,73%	18,57%	61,06%	38,94%	18,21%	67,23%	32,77%	18,94%	70,16%	29,84%	18,88%

6. SONUÇLAR

Sunulan tez çalışması iki ana aşamadan oluşmaktadır; birinci aşamada tutarlılık temelli istatistiksel bir model üzerinde çalışılmış, ikinci aşamada makine öğrenmesi ile istatistiksel modelin geri bildirim saldırıları senaryolarındaki zaafları giderilmeye çalışılmış ve bu amaçla daha zorlu senaryolar ile model sınanmıştır. İstatistiksel tutarlılık modelinin toy hizmet saldırganları senaryolarında neredeyse tamamen koruma sağladığı ve neredeyse bütün hizmet ve geribildirim saldırılarını engelleyebildiği gözlemlenmiştir. Ayrıca hem bireysel hem işbirlikçi iki yüzlü hizmet saldırısı senaryolarında %99,5 seviyelerine kadar ulaşan temiz etkileşim oranları gözlemlenmiştir. Ancak bu başarı oranının yakalandığı %20 saldırgan oranı %40'a çıkarıldığında aynı başarımın sergilenemediği, geri bildirim saldırılarında ise önemli bir koruma/önleme sağlamadığı gözlemlenmiştir. Çalışmamızın ikinci aşamasında istatistiksel modele ek olarak makine öğrenmesi yöntemleri çözüme dahil edilmiştir. Bu aşamada, istatistiksel modelin tespit edemediği saldırıların doğru bir şekilde sınıflandırılması hedeflenmiş, hem de daha karmaşık saldırgan modelleri sınanmış ve uyarlanabilir (*adaptive*) saldırgan davranış senaryolara eklenmiştir. Ayrıca daha karmaşık ve gerçekçi olması için bazı senaryolarda karma saldırgan davranışları deneylere eklenmiştir.

Çizelge 6.1 En iyi sonuçları veren yöntem ve öznitelik kümesi ikilileri

		En Düşük Saldırı Oranı			En Yüksek Doğruluk		
Saldırı Türü	**Senaryo**	**Yöntem**	**Öznitelik Kümesi**	**Sonuç (%)**	**Yöntem**	**Öznitelik Kümesi**	**Sonuç (%)**
Hizmet	Naive Bireysel Saldırgan	Naive Bayes (ML+Tutarlılık)	Tüm Öznitelikler	0,00%	C45 (ML+Tutarlılık)	Tüm Öznitelikler	99,56%
	Hypocritical Bireysel Saldırgan	Isolation Forest (ML+Tutarlılık)	Seçili Öznitelikler	0,00%	Random Forest (ML+Tutarlılık)	Temel+İthal Öznitelikler	98,96%
	Karma Bireysel Saldırgan	Naive Bayes (ML+Tutarlılık)	Seçili Öznitelikler	0,07%	Random Forest (ML+Tutarlılık)	Temel+İthal Öznitelikler	99,67%
	Naive İşbirlikçi Saldırgan	Isolation Forest (ML+Tutarlılık)	Temel+İthal Öznitelikler	0,00%	Random Forest (ML)	Temel+İthal Öznitelikler	99,01%
	Hypocritical İşbirlikçi Saldırgan	Isolation Forest (ML+Tutarlılık)	Seçili Öznitelikler	0,29%	Random Forest (ML+Tutarlılık)	Temel+İthal Öznitelikler	99,24%
Geri Bildirim	Naive Bireysel Saldırgan	Isolation Forest (ML+Tutarlılık)	Tüm Öznitelikler	0,00%	C45 (ML)	Temel+İthal Öznitelikler	99,39%
	Hypocritical Bireysel Saldırgan	Isolation Forest (ML+Tutarlılık)	Seçili Öznitelikler	0,00%	C45 (ML)	Seçili Öznitelikler	99,12%
	Karma Bireysel Saldırgan	Isolation Forest (ML)	Seçili Öznitelikler	0,05%	C45 (ML)	Seçili Öznitelikler	99,21%
	Naive İşbirlikçi Saldırgan	Isolation Forest (ML+Tutarlılık)	Tüm Öznitelikler	0,01%	Random Forest (ML+Tutarlılık)	Temel+İthal Öznitelikler	98,03%
	Hypocritical İşbirlikçi Saldırgan	Naive Bayes (ML+Tutarlılık)	Tüm Öznitelikler	1,26%	Random Forest (ML+Tutarlılık)	Temel+İthal Öznitelikler	91,18%

Çizelge 6.1'de tüm resmi değil yalnızca en iyi sonuçları süzdüğümüz sonuçları görebilirsiniz. Öncelikle sonuçlar bize tüm senaryolarda geçerli tek bir en iyi model olmadığını göstermektedir. Bunun yanı sıra, Isolation Forest yönteminin saldırı engellemede, Random Forest ve C45 yöntemlerinin ise doğru sınıflandırmada diğer yöntemlerden daha iyi olduğu söylenebilir. Ancak Naive Bayes yöntemi, hem saldırı engellemedeki sonuçlarının neredeyse Isolation Forest ile aynı olması, hem de doğruluk ölçütü açısındanyüksek başarım sergilemesinden dolayı genel görünümde daha geniş bir şekilde uygulanabilir yöntemlerden biri olarak gözlemlenmiştir. Ayrıca öznitelik kümesi olarak farklı kümelerin farklı senaryolarda başarılı olduğunu gözlemlememize rağmen, öznitelik seçim algoritmaları ile seçilen öznitelik kümesinin genel çerçevede en iyi ortalama sonuçları verdiğini söyleyebiliriz. Bununla birlikte, en iyi koruma ve en iyi doğruluk başarımlarına neredeyse hiçbir zaman tek bir algoritma, öznitelik kümesi uygulaması ile ulaşılamamıştır. Bunun nedeninin model duyarlılığındaki artışın saldırıları engellemesi ve normal etkileşimlerin saldırı olarak sınıflandırılmasına neden olduğunu düşünebiliriz. Ayrıca buna paralel olarak doğruluğu artırmak istediğimizde genellikle saldırı oranının arttığını gözlemledik.

Isolation Forest, bazı senaryolarda şaşırtıcı derecede başarılı saldırı önleme istatistikleri rapor ederken, bazı senaryolarda sınıflandırmayı tamamen karıştırmaktadır. Tutarlılık temelli filtreleme yapmadığımız senaryolarda da SVM algoritmasının sonuçlarının saldırgan modeline göre büyük ölçüde değişebildiğini gözlemledik. Örneğin, geri bildirim saldırı senaryolarında, karma saldırganları tespit etmede en kötü başarımı gösterirken, toy saldırganları tespit etmenin diğer algoritmalara kıyasla çarpıcı bir gelişme sağladığını gözlemledik.

İşbirliğinin, saldırı modellerinde sonuçları önemli ölçüde etkileyebileceğini gördük. Ayrıca, filtre olarak tutarlılık temelli modeli kullandığımızda modelin işbirlikçi saldırgan senaryolarında genellikle daha iyi sonuçlar verdiğini söylemek için yeterince örnek gördük. Toy saldırganların kolayca tespit edilmesinden ötürü iş birliği yapmaları durumunda saldırı başarısı açısından bir kazanımları olmadığını gözlemledik.

Bazı deney sonuçlarında, özellikle SVM algoritmasının hata oranı ile saldırı oranının eşit ya da çok yakın olduğunu gözlemledik ve ayrıntılı olarak inceledik. Bu tür deney sonuçlarında gördük ki, saldırı oranı ile hata oranı sonuçları yaklaşan deneylerde, yöntemler etkileşim girişimlerini çoğunlukla normal (saldırı değil) olarak sınıflandırmaktadır. Hata oranını artıran her bir sınıflandırma için aynı zamanda saldırıya izin veren model saldırı oranını da artırmış olacaktır. Yani, bu deneylerde uygulanan model, neredeyse hiçbir etkileşim talebini saldırı olarak sınıflandırmayarak saldırıların aynı zamanda hata olarak istatistiklere yansımasına sebep olabilmektedir.

Çalışmamızın bir aşamasında, makine öğrenimi tekniklerini henüz dahil etmediğimiz uygulamalarda, geri bildirim saldırıları temel modelimizde tespit etmekte zorlandığımız saldırı türüydü. Ancak deneylerimizde tıpkı hizmet saldırıları gibi makine öğrenimi ile modelimizi güçlendirerek geri bildirim saldırılarının yüksek oranda tespit edilebildiğini gördük. Ancak seçilen algoritma ve öznitelik kümesinin sonuçlara etkisi yapılan deneyler sonucunda net bir şekilde ortaya konmuştur. Deneylerimiz, tutarlılık temelli modelimiz ile filtrelemek yerine yalnızca makine öğrenmesi yaklaşımının bazı senaryolar için yeterince başarılı olmayabileceğini göstermiştir.

Deneylerin yürütüldüğü ortamdaki gibi bir basitlik, makine öğreniminin daha iyi çalışacağı karmaşık örüntüler için en iyi ortamı sağlamayabilir. Bu nedenle en yalın makine öğrenmesi yöntemlerinden birisi olan Naïve Bayes yönteminin deneylerimizdeki başarısının, bu yalınlığı ile ilişkili olabileceğini değerlendiriyoruz. Makine öğrenmesi kullanıldığında istatistiksel destekleyici yöntemlerle filtrelemenin (tutarlılık temelli modelimizde olduğu gibi) faydalı olduğunu gördük. Özellikle geri bildirim saldırılarında iki yaklaşımın birleştirilmesinin tek başına kullanılmasından çok daha olumlu bir etkisi olduğunu gözlemledik.

Güven yönetimi, literatürde genellikle hizmet sunumu sırasında meydana gelen saldırıları önlemenin sonuçlarına odaklanmaktadır. Önerdiğimiz modelin literatüre katkılarında biri, sadece hizmet saldırılarını değil aynı zamanda geri bildirim saldırılarını da önlemede büyük başarı gösteriyor olmasıdır. Ayrıca modelimizin uygulandığı protokolde ortaya koyduğu bir başka fayda ise; tutarlılığı teşvik eden modelimizin, yüksek tutarlılık

gösteren eşler için yüksek bant genişliği sağlamayı vaat ediyor olmasıdır. Bu şekilde, daha güvenilir eşlerin daha kaliteli hizmet almasına olanak tanıyan daha adil bir ortam sağlamış olur.

Yapılan çalışmalarımız sırasında karşılaştığımız en büyük zorlukların başında, ortak bir veri kümesi kullanarak diğer modeller ile kendi modelimizi karşılaştırma imkansızlığı oldu. Literatüre katkı veren çalışmalar genel olarak güven yönetiminde eşler arası ağların ruhuna uygun olmayan merkezi veri kaynaklarını kullanmayı ya da nasıl bir veri ile çalıştığını paylaşmamayı tercih etmektedir. Bu yüzden literatürde karşılaştırmalı sonuçlara pek rastlayamadığımızı söyleyebiliriz. Bunun yanı sıra, karşılaştırma yapabilmek için başvurduğumuz pek çok yazar program kodlarını paylaşmamış ya da kendi kodları gerçekleştirme çabalarımızda önemli gördüğümüz ancak yayında söz edilmeyen varsayımlar için nasıl karar almamız gerektiğini önermemiştir. Bu açıdan bakıldığında, bu çalışmanın literatüre sunacağı önemli bir diğer katkı da, Peersim üzerinde geliştirilen benzetim ortamı ile sonraki çalışmalarda karşılaştırmayı mümkün kılmak olacaktır.

Bu tez çalışmasında genel olarak varsayım, hizmet olarak sunulan kaynağı belge olduğu varsayılmıştır. Her ne kadar kaynağın cinsinden bağımsız olarak senaryolar kaynak paylaşımı olarak değerlendirilmiş olsa da kaynak türüne, kaynağa özel niteliklere göre öğrenmeyi genişletecek çalışmalar yapılabilir. Bu sayede daha akıllı ve gerçek dünya saldırgan türleri tanımlanabilir ve sistem bu şekilde daha ileri noktalara taşınabilir. Bunun yanı sıra, makine öğrenmesi deneylerimizde, modelin öğrenme aşamasını geleneksel tek seferlik öğrenme (training) ile sınadık. Oysa ki güvenin zaman içinde değişen bir tanımını yapmak da mümkün olabilirdi. Bu gibi bir varsayım evreninde pekiştirmeli öğrenme yöntemleri kullanılarak gerçek zamanlı öğrenme, her etkileşimde öğrenmeye devam etme senaryoları uygulanabilir. Ayrıca derin öğrenme (Deep Learning) ve makine öğrenmesi işini otomatik yerine getiren AutoML (Automated Machine Learning) gibi son yıllarda yaygınlaşan ve araçlara dahil edilen yöntemlerin sonuçlarını gözlemlemek ve karşılaştırmalı olarak paylaşmak literatür için değerli olabilir. Aynı şekilde tahminleme başarımını artırmanın hedeflendiği, birden çok makine öğrenmesi yönteminin tek bir modelde bir araya getirildiği toplu öğrenme (*Ensemble Learning)* yaklaşımı uygulanarak

sonuçlar gözlemlenebilir. Son olarak, bu çalışma kapsamında deney ortamındaki kaynak paylaşımı senaryosu yerine, tutarlılık temelli modelimizi ve çıkardığımız öznitelikleri kullanarak, blok zinciri tabanlı bir eşler arası ağ protokolü üzerinde çalışmalar yapılarak günümüzde parasal bir değere dönüşen bu ağlardaki güven yönetimine katkı verilebilir.

7. KAYNAKLAR

[1] M. Nofer, P. Gomber, O. Hinz ve D. Schiereck, «Blockchain,» *Business & Information Systems Engineering,* cilt 59, no. 3, p. 183–187, 2017.

[2] H. Wang, Z. Zheng, S. Xie, H.-N. Dair, Chen ve X. Chen, «Blockchain challenges and opportunities: a survey,» *International Journal of Web and Grid Services,* cilt 14, no. 4, pp. 352-375, 2018.

[3] M. Halis, A. Şenkal ve O. Türkay, Türk İşletmelerinde Ortaklık ve Güven, İstanbul: İstanbul Ticaret Odası, 2009, p. 49.

[4] J. Benhabib, A. Bisin ve Matthew O. Jackson, Handbook of social economics., Elsevier, 2010.

[5] A. Montresor and M. Jelasity, "A scalable P2P simulator:PeerSim," in *IEEE International Conference on Peer-to-Peer Computing*, 2009.

[6] S. R. Garner, «Weka: The waikato environment for knowledge analysis,» *Proceedings of the New Zealand computer science research students conference*, 1995.

[7] R. Rathinasamy ve L. Raj, «Comparative Analysis of C4.5 and C5.0 Algorithms on Crop Pest Data,» *International Journal of Innovative Research in Computer and Communication Engineering,* cilt 5, no. 1, pp. 50-58, 2019.

[8] «Support-vector machine,» 20 06 2022. [Çevrimiçi]. Available: https://en.wikipedia.org/wiki/Support-vector_machine. [Erişildi: 01 07 2022].

[9] A. Krishnan, «Anomaly Detection with Isolation Forest & Visualization,» [Çevrimiçi]. Available: https://towardsdatascience.com/anomaly-detection-with-isolation-forest-visualization-23cd75c281e2. [Erişildi: 01 07 2022].

[10] G. Biau ve E. Scornet, «A random forest guided tour,» *TEST,* cilt 25, p. 197–227 , 2016.

[11] «Random Forest,» [Çevrimiçi]. Available: https://en.wikipedia.org/wiki/Random_forest. [Erişildi: 01 07 2022].

[12] A. Kumar, «Naive Bayes Classifier: Calculation of Prior, Likelihood, Evidence & Posterior,» [Çevrimiçi]. Available: https://medium.com/@abhishek.km23/naive-bayes-classifier-calculation-of-prior-likelihood-evidence-posterior-74d7d27eec24. [Erişildi: 01 07 2022].

[13] S. Çimen ve B. Bolat, «Diagnosis of Parkinson's Disease by using ANN,» *International Conference on Global Trends in Signal Processing, Information Computing and Communication*, 2016.

[14] M. Jelasity, S. Voulgaris, R. Guerraoui, A.-M. Kermarrec ve M. v. Steen, «Gossip-based peer sampling,» *ACM Transactions on Computer Systems,* cilt 25, no. 3, 2007.

[15] K. Aberer and Z. Despotovic, "Managing Trust in a Peer-2-Peer Information System," *Proc. 10th Int'l Conf. Information and Knowledge Management (CIKM),* 2001.

[16] S. D. Kamva, M. T. Schlosse and H. G. Molina, "The Eigentrust Algorithm for Reputation Management in P2P Networks," in *12th World Wide Web Conference*, 2003.

[17] X. Li and L. Liu, "Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities," *IEEE Transactions on Knowledge and Data Engineering,* vol. 16, no. 7, pp. 843-857, 2004.

[18] G. Wang, F. Musau, S. Guo and M. B. Abdullahi, "Neighbor Similarity Trust against Sybil Attack in P2P E-Commerce," *IEEE Transactions on Parallel and Distributed Systems,* vol. 26, no. 3, pp. 824-833, 2015.

[19] L. Guo, S. Yang, J. Wang and J. Zhou, “Trust model based on similarity measure of vectors in P2P networks,” in *International Conference on Grid and Cooperative Computing*, 2005.

[20] Y.-m. Liu, S.-b. Yang and L.-t. Guo, “A Distributed Trust-based Reputation Model in P2P System,” in *Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 2007.

[21] J. Hu, Q. Wu and B. Zhou, “FCTrust: a robust and efficient feedback credibility-based distributed P2P trust model,” in *IEEE Young Computer Scientists*, 2008.

[22] A. B. Can and B. Bhargava, “Sort: A self-organizing trust model for peer-to-peer systems,” *IEEE Transactions on dependable and secure computing,* vol. 10, no. 1, pp. 14-27, 2013.

[23] F. Cornelli, . E. Damiani, S. D. C. di Vimercati, S. Paraboschi and P. Samarati, “Choosing Reputable Servents in a P2P Network,” in *11th World Wide Web Conf. (WWW)*, 2002.

[24] A. A. Selcuk, E. Uzun and M. R. Pari, “A reputation-based trust management system for P2P networks.,” in *IEEE Cluster Computing and the Grid*, 2004.

[25] R. Zhou, K. Hwang and M. Cai, “GossipTrust for Fast Reputation Aggregation in Peer-to-Peer Networks,” *IEEE Transactions on Knowledge and Data Engineering,* vol. 20, no. 9, pp. 1282-1295, 2008.

[26] Z. Su, L. Liu, M. Li, X. Fan and Y. Zhou, “ServiceTrust: trust management in service provision networks,” in *Services Computing (SCC)*, 2013.

[27] Z. Su, L. Liu, M. Li, X. Fan and Y. Zhou, “Reliable and Resilient Trust Management in Distributed Service Provision Networks,” *CM Transactions on the Web (TWEB),* vol. 9, no. 3, 2015.

[28] J. V. Yao Wang, «Bayesian Network Trust Model in Peer-to-Peer Networks,» *Agents and Peer-to-Peer Computing*, Berlin, 2004.

[29] G. Swamynathan, B. Y. Zhao ve K. C. Almeroth, «Decoupling Service and Feedback Trust in a Peer-to-Peer Reputation System,» *Parallel and Distributed Processing and Applications - ISPA 2005 Workshops*, Nanjing, 2005.

[30] M. Chen, K. Kita ve X. Luo, «Cluster-Based Reputation Model in Peer-to-Peer Network,» *International Journal of Machine Learning and Computing,* cilt 1, no. 4, pp. 366-371, 2011.

[31] Q. Han, H. Wen, M. Ren, B. Wu ve S. Li, «A topological potential weighted community-based recommendation trust model for P2P networks,» *Peer-to-Peer Networking and Applications,* cilt 14, no. 3, p. 1048–1058, 2015.

[32] E. Zhai, H. Sun, S. Qing ve Z. Chen, «Sorcery: Overcoming deceptive votes in P2P content sharing systems,» *Peer-to-Peer Networking and Applications,* cilt 4, pp. 178-191, 2011.

[33] K. Chen, G. Liu, H. Shen and F. Qi, "Sociallink: utilizing social network and transaction links for effective trust management in P2P file sharing systems," in *IEEE International Conference on Peer-to-Peer Computing (P2P)*, Boston, 2015.

[34] Z. Xei, Y. Geng and J. Bi, "STTM: Similarity Transitivity Chain Based Trust Model in P2P Environment," in *Communications (ICC), 2010 IEEE International Conference*, 2010.

[35] A. Das ve M. M. Islam, «A Novel Feedback Based Fast Adaptive Trust Model for P2P Networks,» *IEEE Local Computer Network Conference*, 2010.

[36] R. Prasad, V. Srinivas, V. Kumari ve K. Raju, «An Effective Calculation of Reputation in P2P Networks,» *Journal of Networks,* cilt 4, no. 5, pp. 332-342, 2009.

[37] K. X. Ouyang, B. Vaidya ve D. Makrakis, «A Probabilistic-Based Trust Evaluation Model Using Hidden Markov Models and Bonus Malus Systems,» *IEEE Third International Conference on Social Computing*, 2011.

[38] S. Z. Hamdi Yahyaoui, «Bootstrapping trust of Web services based on trust patterns and Hidden Markov Models,» *Knowledge and Information Systems,* cilt 37, pp. 389-416, 2013.

[39] G. Huang, M. Hu, Y. Zhou, P. Liu ve Y. Zhang, «A Distributed Trust Model Based on Reputation Management of Peers for P2P VoD Services,» *KSII Transactions on Internet and Information Systems,* cilt 6, no. 9, pp. 2285-2301, 2012.

[40] B. Y. Chunqi Tian, «A D-S evidence theory based fuzzy trust model in file-sharing P2P networks,» cilt 7, pp. 332-345, 2014.

[41] L. Guo, Y. Luo, Z. Zhou and M. Ji, "A recommendation trust method based on fuzzy clustering in P2P networks," *Journal of Software,* vol. 8, no. 2, pp. 357-360, 2013.

[42] J. Saeed, M. Shojafar, S. Shariatmadari and S. S. Ahrabi, "FR trust: a fuzzy reputation–based model for trust management in semantic P2P grids," *International Journal of Grid and Utility Computing,* vol. 6, no. 1, pp. 57-66, 2014.

[43] R. Feng, S. Che, X. Wang ve N. Yu, «Trust Management Scheme Based on D-S Evidence Theory for Wireless Sensor Networks,» *International Journal of Distributed Sensor Networks,* 2013.

[44] S. Chen, G. Wang ve W. Jia, «Cluster-group based trusted computing for mobile social networks using implicit social behavioral graph,» *Future Generation Computer Systems,* cilt 55, pp. 391-400, 2016.

[45] W. Yuan, D. Guan, S. Lee ve Y. Lee, «A Dynamic Trust Model Based on Naive Bayes Classifier for Ubiquitous Environments,» *High Performance Computing and Communications*, Munich, 2006.

[46] G. D'Angelo, S. Rampone ve F. Palmieri, «Developing a trust model for pervasive computing based on Apriori association rules learning and Bayesian classification,» *Soft Computing,* cilt 21, p. 6297–6315, 2017.

[47] H. Baohua, H. Heping ve L. Zhengding, «Identifying local trust value with neural network in P2P environment,» *IEEE and IFIP International Conference in Central Asia on Internet*, 2005.

[48] X. Liu, G. Tredan ve A. Datta, «A generic trusted framework for large-scale open systems using machine learning,» *Computational Intelligence,* cilt 30, no. 4, pp. 700-721, 2013.

[49] A. Aref ve T. Tran, «A hybrid trust model using reinforcement learning and fuzzy logic,» *Computational Intelligence,* cilt 34, no. 2, pp. 515-541, 2017.

[50] H. Yahyaoui ve A. Al-Mutairi, «A feature-based trust sequence classification algorithm,» *Information Sciences,* cilt 328, pp. 455-484, 2016.

[51] N. Korovaiko ve A. Thomo, «Trust prediction from user-item ratings,» *Social Network Analysis and Mining,* cilt 3, pp. 749-759, 2013.

[52] C. Mao, R. Lin, C. Xu ve Q. He, «Towards a Trust Prediction Framework for Cloud Services Based on PSO-Driven Neural Network,» *IEEE Access,* cilt 5, pp. 2187-2199, 2017.

[53] U. Jayasinghe, G. M. Lee, T.-W. Um ve Q. Shi, «Machine Learning Based Trust Computational Model for IoT Services,» *IEEE Transactions on Sustainable Computing,* cilt 4, no. 1, pp. 39-52, 2019.

[54] R. Mohanty, V. Ravi ve M. R. Patra, «Web-services classification using intelligent techniques,» *Expert Systems with Applications,* cilt 37, no. 7, pp. 5484-5490, 2010.

[55] H. Liu, E.-P. Lim, H. Lauw, M.-T. Le, A. Sun, J. Srivastava ve Y. Kim, «Predicting Trusts among Users of Online Communities – an Epinions Case Study,» *Proceedings of the 9th ACM conference on Electronic commerce*, 2008.

[56] K. Zolfaghar ve A. Aghaie, «A syntactical approach for interpersonal trust prediction in social web applications: Combining contextual and structural data,» *Knowledge-Based Systems,* cilt 26, pp. 93-102, 2012.

[57] W. Yuji, «The Trust Value Calculating for Social Network Based on Machine Learning,» *9th International Conference on Intelligent Human-Machine Systems and Cybernetics*, 2017.

[58] A. Papaoikonomou, M. Kardara ve T. Varvarigou, «Trust Inference in Online Social Networks,» *International Conference on Advances in Social Networks Analysis and Mining*, 2015.

[59] X. Chen, Y. Yuan ve M. A. Orgun, «Using Bayesian networks with hidden variables for identifying trustworthy users in social networks,» *Journal of Information Science,* cilt 46, no. 5, pp. 600-615, 2020.

[60] Y. A. Kim ve H. S. Song, «Strategies for predicting local trust based on trust propagation in social networks,» *Knowledge-Based Systems,* cilt 24, no. 8, pp. 1360-1371, 2011.

[61] R. V. Barenji, «A blockchain technology based trust system for cloud manufacturing,» *Journal of Intelligent Manufacturing ,* 2021.

[62] X. Chen, J. Ding ve Z. Lu, «A decentralized trust management system for intelligent transportation environments,» *IEEE Transactions on Intelligent Transportation Systems,,* cilt 23, no. 1, pp. 558-571, 2022.

[63] W. Abdelghani, I. Amous, C. A. Zayani, F. Sèdes ve G. Roman-Jimenez, «Dynamic and scalable multi-level trust management model for Social Internet of Things,» *The Journal of Supercomputing,* cilt 78, p. 8137–8193, 2022.

[64] J. L. Xu Wu, «A blockchain-based trust management method for Internet of Things,» *Pervasive and Mobile Computing,* cilt 72, 2021.

[65] A. Das and M. M. Islam, "SecuredTrust: A Dynamic Trust Computation Model for Secured Communication in Multiagent Systems," *IEEE Transaction on Dependable and Secure Computing,* vol. 9, no. 2, pp. 261 - 274, 2012.

[66] L. M. Aiello ve G. Ruffo, «Secure and Flexible Framework for Decentralized Social Network Services,» *8th IEEE International Conference on Pervasive Computing and Communications Workshops*, 2010.

[67] H. Li ve M. Singhal, «Trust management in distributed systems,» *Computer ,* cilt 40, no. 1, pp. 45-53, 2007.

[68] L. N. Jianyu Miao, «A Survey on Feature Selection,» *Procedia Computer Science,* cilt 91, pp. 919-926, 2016.

[69] W. Javadoc, «CorrelationAttributeEval,» [Çevrimiçi]. Available: https://weka.sourceforge.io/doc.dev/weka/attributeSelection/CorrelationAttributeEval.html. [Erişildi: 01 07 2022].

Tezden Türetilmiş Yayınlar

Yasin Şahin ve Ahmet Burak Can , "Consistency-based trust management in P2P networks", *Turkish Journal of Electrical Engineering and Computer Science*, c. 26, sayı. 2, ss. 631-643, Nis. 2018

.