



Heterojen Bilgi Ağlarında Zamana Duyarlı Bilimsel Takım Önerme

Nuri ÖZDEMİR*,^a Hadi GÖKÇEN

^{a,*} Gazi Üniversitesi, Biliřim Enstitüsü, Yönetim Biliřim Sistemleri Bölümü, ANKARA, 06680, TÜRKİYE

^b Gazi Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendislięi Bölümü, ANKARA, 06560, TÜRKİYE

MAKALE BİLGİSİ

Alınma: 19.03.2026
Kabul: 08.04.2026

Anahtar Kelimeler: takım oluřturma, heterojen bilgi ağları, sosyal ağ analizi, ağ veritabanı, kurumsal baęlantı, neo4j

*Sorumlu Yazar

e-posta:
nuri.ozdemir1@gazi.edu.tr,
nuri.ozdemir@tubitak.gov.tr

ÖZET

Geniř kapsamlı disiplinlerarası bilimsel arařtırma projelerinin artması, farklı uzmanlıklara sahip arařtırmacıların proje takımlarında bir araya getirilmesini gerektirmektedir. Literatürde sosyal ağlara veya heterojen bilgi ağlarına (HBA) dayalı olarak takım üretmeyi hedefleyen çalışmalar olmakla birlikte, mevcut çalışmalar çoęunlukla ortak yazarlık veya kiři-uzmanlık/yayın yeri iliřkilerine dayanmakta, aynı kurumda çalışma iliřkilerini göz ardı etmektedir. Bu çalışmada, YÖK Akademik Platformu'ndan alınan bilgiler hem ortak yazarlık hem de aynı kurumda çalışma iliřkilerini kapsayan bir heterojen bilgi ağına dönüřtürülmüř ve bu ağ üzerindeki baęlantıları üç aşamalı bir stratejiyle kullanarak, çalışmaları güncel, uzmanlık derecesi yüksek, iletiřim maliyeti düřük takımlar oluřturan, zamana duyarlı bir algoritma (ZDA) önerilmiřtir. Bu algoritmanın, ortak kurumda geirilen zamanın akıřmasını dikkate alan versiyonun (ZDA-3 Yıl) yanı sıra, bu akıřmayı kontrol etmeyen (ZDA-Zamansız) versiyonu ve yalnızca ortak yazarlık iliřkilerine dayanan (OYDA) versiyonu da geliřtirilmiřtir. 200 farklı anahtar kelime listesiyle yapılan testlerde ZDA algoritmaları %94,5 tamamlanma oranına ulařırken, daha geniř arama mesafesine sahip OYDA algoritması %88'de kalmıřtır. Sonuçlar, HBA'daki ok katmanlı iliřki yapısının takım oluřurmada etkili olduęunu, aynı kurumda çalışma iliřkilerinin hem tamamlanma oranını artırdıęını hem de daha düřük maliyetli takımlar oluřurmaya imkân verdięini ortaya koymaktadır.

DOI: 10.59940/jismar.1912778

Time Sensitive Scientific Team Recommendation on Heterogeneous Information Network

ARTICLE INFO

Received: 19.03.2026
Accepted: 08.04.2026

Keywords: team formation, heterogeneous information network, social network analysis, graph database, institutional affiliation, neo4j

*Corresponding Authors

e-mail:
nuri.ozdemir1@gazi.edu.tr,
nuri.ozdemir@tubitak.gov.tr

ABSTRACT

The increasing number of large-scale interdisciplinary scientific research projects requires researchers with diverse expertise to come together in project teams. While there are studies in the literature that aim to suggest teams based on social networks or Heterogeneous Information Networks (HIN), existing studies mostly rely on co-authorship or author-skill/venue relationships and disregard institutional affiliations. In this study, the data obtained from the YÖK Academic Platform is transformed into a HIN encompassing both co-authorship and institutional affiliations. A time-sensitive algorithm (TSA) is proposed on this network, using a three-stage strategy employing co-authorship, institutional affiliation, and remote search to create teams with up-to-date work, high levels of expertise, and low communication costs. In addition to a version that considers the overlap of time spent at the same institution in institutional affiliations (TSA-3 Years), versions that do not control for this overlap (TSA-Timeless) and that rely solely on co-authorship relationships (COAT) have also been developed. In tests conducted with 200 different keyword lists, TSA algorithms achieved a 94.5% completion rate, while COAT, despite having a wider search distance, remained at 88%. The results reveal that the multi-layered relational structure in HINs is effective in team formation, and that institutional affiliation relationships both increase the completion rate and enable the creation of lower-cost teams.

DOI: 10.59940/jismar.1912778

1. GİRİŞ (INTRODUCTION)

Projeler karmaşıktıkça onları tek bir kişinin gerçekleştirmesi zorlaşır ve başarıyla tamamlamak için projede gerekli yetkinliklere sahip farklı kişilerle bir takım oluşturmak gerekir. Proje sponsorunun veya yöneticisinin kişisel değerlendirmeleriyle bir takım oluşturmak mümkün olsa da, birbiriyle uyumlu kişilerden oluşan yetkin ve yüksek performanslı bir takım oluşturmak için analitik araçların kullanılması gerektiği açıktır [1]. Literatürde takım oluşturma için önerilen ilk yöntemler [2, 3] büyük ölçüde kişilerin yetkinliklerinin takımda ihtiyaç duyulan becerilerin eşleştirilmesi ile sınırlı kalmış, bir araya gelecek kişiler arasındaki iletişim, uyumlu çalışma kapasitesi vb. birçok etken bu çalışmalarda göz ardı edilmiştir. [4-8] kapsamında takım oluşturmaya sosyal etkenleri de dahil etmek gerektiği belirtilmiş, çeşitli öneriler de sunulmuş ancak takımı ve takımın oluşturacak aday havuzunu bir sosyal ağ olarak modelleyen analitik bir yöntem ilk kez Lappas vd. [9] tarafından ortaya konmuştur. Söz konusu çalışmadaki, yazarlar arası ortak yazarlık ilişkilerine dayalı homojen sosyal ağda, gerekli becerileri sağlarken iletişim maliyetini de enküçükleme yaklaşımı daha sonraki pek çok çalışmaya da yol göstermiştir. [10-21] gibi pek çok takım oluşturma çalışmasında bir veya birden fazla görev için takımlar oluşturulurken benzer yöntemlerle bir homojen ağ oluşturulmuş, bu ağdaki kişilerle ve ağın genel yapısı ile ilgili kısıtlar, iletişim maliyetini en küçükmeye çalışılırken göz önünde bulundurulmuştur. [36, 37, 39] gibi çalışmalarda ise literatüre Sun ve Han tarafından kazandırılan heterojen bilgi ağı [33] kavramı, takım oluşturma problemi ile birleştirilmiştir. Böylece birden fazla düğüm/kenar türü barındıran heterojen bir ağ kullanarak daha gerçekçi ve verimli takımlar üretilebilmek mümkün olmuştur.

[9-22] sezgisel veya meta-sezgisel yöntemlere dayanmakla birlikte güncel bazı çalışmalarda takım oluşturma [38-40], mevcut takım performansını iyileştirme/koruma [28], ağda uzman/işbirlikçi bulma [44-45] vb. amaçlarla makine öğrenmesine dayalı yöntemler de kullanılmaya başlamıştır. Bu tür yöntemler son zamanlarda farklı çalışmalarda popüler olsa da, eğitim maliyeti, sonuçların öngörülebilirliğinin ve açıklanabilirliğinin zayıf olması ve birçok zaman üretilen takımda yazarlar arası bağlantılılık gibi sıkı koşulların sağlanmaması vb. eksiklikleri vardır. Literatürdeki, sosyal ağlara ya da heterojen bilgi ağlarına dayalı sezgisel yaklaşımlarda ise, takım oluşturmada etkisi kritik olabilecek, aynı kurumda çalışma ilişkisinin (kurumsal bağlantı) analitik yöneme dahil edilmediği görülmüştür. Bunun

yapılmamasındaki en önemli neden, kişilerin kurumlarla ilişkilerinin sistemli ve doğru bir şekilde kaydedildiği verilere ulaşmadaki zorluktur.

Bu çalışma kapsamında YÖK Akademik Platformu'ndan alınan, Türkiye'deki üniversitelerde görevli, Endüstri Mühendisliği veya Yönetim Bilişim Sistemleri bilim alanında çalışan iki binin üzerindeki öğretim elemanı, bunların çalıştıkları üniversite/fakülte/bölümler, bilimsel yayınları ve bu yayınlardaki ortak yazarlarını da içeren bir heterojen bilgi ağı oluşturulmuştur. Platformda öğretim elemanlarının görev yeri bilgileri, başlangıç ve bitiş tarihleri ile birlikte, düzenli bir biçimde tutulduğu için bu verilerle kişiler arası aynı kurumda çalışma ilişkilerini zamana duyarlı olarak kaydetmek ve kullanmak mümkündür. Ayrıca, kişilerin makalelerinin anahtar kelimelerinden alınan uzmanlık alanlarını da makale yılları ile birlikte göz önünde bulundurarak, ihtiyaç duyulan alandaki uzmanlığın güncel olup olmadığını da dikkate almak mümkündür.

Çalışmamızda, zaman bilgileri ile birlikte kaydedilen ortak yazarlık ve aynı kurumda çalışma bağlantıları kullanılarak aşağıdaki sorulara yanıtlar verilmesi hedeflenmiştir:

1. Aynı kurumda çalışma ilişkilerini göz önünde bulundurmanın, gerekli tüm uzmanlıkları barındıran, tamamlanmış takımlar üretmede etkisi nedir?
2. Kişilerin (öğretim elemanı), ilgili kurumda çalıştığı yılların çalışma durumunu dikkate almanın, gerekli tüm uzmanlıkları barındıran, tam takımlar üretmede etkisi nedir?

Özetle, bu çalışmada, öğretim elemanlarının makaleleri ile birlikte, çalıştıkları kurumları ve ilgili kurumlarda çalışma tarihlerini içeren heterojen bir bilgi ağı girdi olarak verilecek, takımların tamamlanma oranı, iletişim maliyeti ve yetkinlik düzeyi çıktı olarak elde edilecektir. Aynı kurumda çalışma ilişkisi ve bu ilişkinin tarihsel çakışması bağımsız değişkenler, takımların tamamlanma oranı, iletişim maliyeti ve yetkinlik düzeyi bağımlı değişkenler olarak kullanılacaktır.

2. LİTERATÜR İNCELEMESİ (LITERATURE REVIEW)

Takım oluşturmak amacıyla analitik bir yöntem öneren ilk çalışmanın Zakarian ve Kusiak [2] tarafından gerçekleştirildiği söylenebilir. Baykasoglu vd. [3] de bu problem için bulanık derecelendirme ve

tavlama benzetimi modeline dayanan bir analitik yöntem önermişlerdir. Ancak bu çalışmalarda, projede ihtiyaç duyulan tüm yetkinlikleri taşıyan kişilerin belirlenmesine odaklanılmış, takımın üyeleri arasındaki iletişim maliyeti hesaba katılmamıştır. [4-8] numaralı çalışmalarda yazarların aslında birbirleriyle bağlantılı olduğundan ve bir ağ meydana getirdiğinden bahsedilse de, takımı oluşturacak aday havuzunu ve oluşan takımı bir sosyal ağ olarak modelleyen analitik bir yöntem önerilmemiştir.

Lappas [9], takım oluşturma problemi literatüründe çığır açıcı olan çalışmasında, bilimsel yayın yazarlarını düğüm/nokta (node), ortak yazarlık ilişkilerini ise kenar/bağlantı (edge) olarak modelleyen ağırlıklı ve yönsüz bir sosyal ağ tanımlamıştır. Aday havuzunu temsil eden bu ağda, daha önce ortak yayın yapmış kişiler birbirlerine tek bir çizgi ile bağlıdır. Ortak yazarın ortak yazarına erişmek için ise iki adım gerektirir. Bağlantılara ağırlık atanmışsa mesafeler sadece adım sayısı olarak değil, noktaları birbirine bağlayan en kısa yoldaki kenarların ağırlıkları toplanarak hesaplanır. Tüm bağlantılar birlikte düşünüldüğünde birbirine doğrudan veya dolaylı olarak bağlı yazarlardan oluşan bir sosyal ağ meydana gelir. [9], oluşacak takımın, büyük ağın bir alt kümesi olduğunu, takımda gerekli uzmanlıkları sağlamanın yanında seçilen kişiler arası mesafelerin kısa olmasıyla, takım içi iletişim maliyetinin azalacağını ve dolayısıyla takım performansının artacağını öne sürmüştür. Bu doğrultuda, gerekli yetkinlikleri sağlamayı mutlak kısıt olarak kullanan, takım üyeleri arasındaki iletişim maliyetini de enküçükleyen bir alt çizge (subgraph) bulma problemi tanımlanmış ve sezgisel yöntemlerle probleme çözüm bulmuştur. Daha sonra, sosyal ağda takım oluşturma problemi, takımda ilgili yetkinliklere sahip olması gereken kişi sayısı da bir değişken olarak tanımlanarak genelleştirilmiştir [10]. Lappas vd. çalışmasında [4] “var” veya “yok” kabul edilen yetkinliğin, takımda ihtiyaç duyulan ve yazarlarda bulunan seviyeleri, Dorn and Dustdar [11] tarafından hesaba katılmıştır. Kargar ve An [12] ise en yetkin k tane takımı önerecek şekilde amaç fonksiyonunu güncellemiştir.

Anagnostopoulos vd., bir akış olarak gelen görevlerin birden fazla takıma adil atanmasına odaklanmış [13] daha sonraki çalışmalarında ise bu amaca iletişim maliyetinin enküçüklenmesini de eklemişlerdir [14]. Majumder vd. [15] bir kişinin birden fazla göreve atanabilmesi ile ilgili sınır getiren iş yükü ve kapasite değişkenlerini dahil etmiştir. Uzak mesafedeki kişilerin takıma alınmasını engelleyen eşik uzaklık değeri [16] ve kişiler arası güven değeri [17] gibi farklı değişkenler de takım oluşturma problemine dahil edilmiştir.

Literatürdeki takım oluşturma çalışmaları genellikle birbirleriyle çelişkili olan performansla veya takım yapısıyla ilgili birden fazla hedefi dengelemeye çalışmıştır. Örneğin [18] çevrimiçi sosyal ağlarda takım oluşturmaya odaklanan çalışma kapsamında, takım performansının yanında, üyelerin mahremiyetle ilgili çekincelerini de gözetecek şekilde takım oluşturmaya önermiştir. [19] iletişim maliyetini, uzmanlık düzeyini, müşterek güven puanı, profil benzerliği ve coğrafi yakınlık değerlerini takım oluşturma probleminde bir arada kullanmıştır. [20] ağdaki farklı kliklerden üyeler alarak takımda çeşitliliği artırmaya çalışmış, böylece öneri değerlendirme gibi amaçlara hizmet edebilecek, yanlılığın azaltıldığı takımlar üretmeye odaklanmıştır.

Gutiérrez vd. [21] bireylerin zamanlarını birden fazla proje arasında bölüştürebildiği Çoklu Takım Oluşturma Problemini ortaya koymuş; bu çalışmada da kişiler arası ilişkiler modele dahil edilmiştir. [22] bir havuzdan optimal takımı en başta üretmek yerine, sezgisel bir yöntemle aday takımlar üretmiş ve tam sayılı programlama kullanarak aday takımlar arasından en uygununu seçmiştir. [23] de Çoklu Takım Oluşturma Problemi kapsamında tamsayı programlamayı kullanmıştır.

[24] takımdan üyelerin ayrılması riskini modellemek için güvenilirliği bir parametre olarak öne sürmüş ve takımlar için ana ve yedek üyeler önermiştir. IDEA Lab grubu [25-28], sıfırdan takım oluşturmak yerine mevcut takımları korumak veya iyileştirmek için (değiştirme, iyileştirme, genişleme, küçülme gibi ihtiyaçlara yanıt olarak) ağ tabanlı yaklaşımlar kullanmıştır. [25] ve [27] ağ tabanlı benzerliğe dayalı olarak takımın yeniden yapılandırılmasını araştırmış, [28] dinamik takım optimizasyonu için pekiştirmeli öğrenme uygulamış, [26] ise takımın bir bölümünü (sub-team) değiştirmede doğruluk ve verimliliği artırmak için kümeleme tabanlı bir Grafik Sinir Ağı çerçevesi kullanmıştır.

Son zamanlarda yapılan bazı çalışmalar [29-32] uzmanları sosyal ağdaki düğümler olarak temsil etmiş, ancak ilişkileri, dolayısıyla iletişim maliyetini ortak yazarlık geçmişi yerine yetkinlik örtüşmesinden türetmiştir; yani bu çalışmalarda, iki kişinin ortak yetkinlikleri varsa bu iki kişi arasında bir bağ olduğu kabul edilmiştir. İletişim maliyetini en aza indirmek için [29], yeni bir takas operatörü kullanan, geliştirilmiş bir parçacık sürü optimizasyon algoritması önermiştir; [30] ise Jaya algoritmasını değiştirilmiş bir takas operatörüyle uyarlamıştır. Rehman vd. [32] ekip oluşturma probleminin yüksek zaman karmaşıklığının üstesinden gelmek için, iletişim maliyetlerini hesaplamadan önce ağı yalnızca

ilgili uzmanlara ve becerilere indirgeyerek bir arama alanı indirgeme stratejisi kullanmıştır.

Literatürdeki birçok takım oluşturma çalışmasının ortak özelliği, ortak yazarlığa veya ortak yetkinliklere dayalı bir ilişki tanımı ile, sadece yazarlar arası ilişkileri modelleyen homojen bir sosyal ağa dayanmasıdır. Ancak, takım oluşturmada adaylar arası ilişkileri daha iyi modellemek ve daha gerçekçi çözüm önerileri sunabilmek için, düğüm ve kenarların indirgendiği homojen sosyal ağlar yerine Sun ve Han [33] tarafından literatüre kazandırılan Heterojen Bilgi Ağlarını (HBA) kullanmak daha etkili bir tercih olarak ortaya çıkmıştır. Literatürde çok ilişkili ağlar [34] veya çoklu ağlar [35] olarak da adlandırılan bu ağ yapısı ile, birden fazla nokta/düğüm türü ve birden fazla ilişki türünü modelleyerek akademik yapıların daha doğru, kapsamlı temsillerini oluşturmak mümkündür.

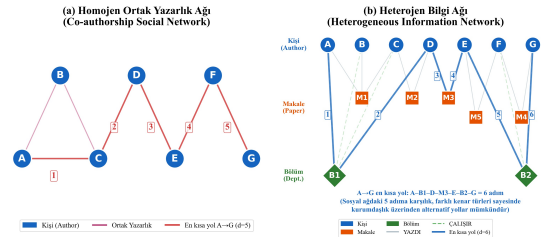
Örneğin, [36] “beceri” düğümleri ile “yazar” düğümleri arasındaki ilişkileri içeren heterojen bir ağda ekip oluşturmaya formüle etmiş, becerileri düğüm olarak tanımlamanın hesaplama verimliliğini artırdığını savunmuştur. Bu çalışma, takım oluşturmada kişilere ve yetkinliklerine odaklandığı için literatürdeki önceki çalışmalardan kapsam olarak çok ayrılmamıştır. Ancak, [37] ek düğüm türlerini (örneğin, yayın yeri, konu ve makale) ve ilişki türlerini (örneğin, yazarlık ve alıntı) dahil etmiş; takım içi iletişim maliyetini yazar düğümlerini birbirine bağlayan meta-yollara dayalı olarak hesaplayarak ekip oluşumu için heterojen bilgi ağında, kapsamlı bir çerçevede ele almıştır.

Takım oluşturma ile ilgili güncel çalışmalarda homojen veya heterojen ağlarda uzman ve beceri/yetenek temsillerini öğrenmek için makine öğrenmesi ve gömme (embedding) tekniklerinden yararlanılmaya başlanmıştır. Örneğin, [38] mevcut çalışmalardaki seyrek veri ve ölçeklenebilirlik sorunlarını dikkate alarak, gerekli beceriler ve uzmanlar arasındaki eşlemeleri öğrenmek için Varyasyonel Bayes Sinir Ağlarını (VBNN'ler) kullanmıştır. [39] heterojen ağlar içindeki geçmiş işbirliği modellerini dikkate alarak uzmanların ve becerilerin temsiliyi iyileştirmek için aktarım öğrenmesini uygulamıştır. [40] yazarlar, makaleler, mekanlar ve beceriler gibi farklı düğüm türleri arasındaki hem yerel hem de küresel ilişkileri yakalayan bir alt-çizge temsil öğrenmesi yaklaşımı önermiştir. [41] ise gömme temelli öğrenmeye dayanan çalışmada, araştırmacıların kurumlarının ilgi alanlarını modele dahil ederek kurumlar arası takımların kurulmasını kolaylaştırmıştır.

[42] homojen bir ağdan dengeli takımlar oluşturmak için Graf Evrimsel Ağları (Graph Convolutional Networks, GCN) ve Karma Tamsayı Doğrusal Programlamayı (MILP) birleştiren hibrit bir model önermiştir. [43] gerekli becerileri sağlarken takım içi işbirliği uyumunu da optimize etmek amacıyla, geçmiş takım verilerinden yararlanan Geri Getirme Destekli Üretim (Retrieval Augmented Generation: RAG) tabanlı üretken yapay zeka modeli geliştirmiştir.

Takım oluşturma çalışmalarına benzer şekilde, bilimsel çalışmalar için işbirlikçi/ortak veya uzman tavsiyesi sunan bazı çalışmalarda da [44-46] benzer ağ temsilleri ve makine öğrenmesi yöntemleri kullanılmıştır. [47] ise heterojen bilgi ağı kullanmasa da çoklu ilişki türlerini barındıran heterojen bir hiper-çizge kullanarak, işbirlikçi tavsiyesi çalışmalarında ilişkilerin heterojenliğinin modellenmesinin önemini göstermiştir. [48] uzman/takım tavsiyesine odaklanmasa da makaleler, patentler, yazarlar ve mucitlerden oluşan heterojen bir ağı analiz ederek, belirli yapısal ve ilişkisel konfigürasyonların çalışmalardan ortaya çıkacak özgünlük ve etkiyle güçlü bir şekilde ilişkili olduğunu göstermiştir.

Takım oluşturma problemini yalnızca ortak makale yazarlığı ilişkisine dayanan bir sosyal ağ üzerinde veya bizim kurguladığımız şekilde, kurum ilişkilerini de kapsayan bir heterojen bilgi ağı üzerinde modellemenin olası etkilerini Şekil 1’de görmek mümkündür.



Şekil 1. Sosyal Ağ ve Heterojen Bilgi Ağı Karşılaştırması

(Social Network vs Heterogeneous Information Network Comparison)

Şekil 1’de gösterilen 7 kişilik aday havuzunda, yetkinlikleri nedeniyle A ve G kişileri ile 2 kişilik bir takım oluşturacağımızı varsayalım. Sadece ortak yazarlık ilişkilerine dayanan Şekil 1-a’da A ve G kişilerini birbirine bağlayan yol uzunluğu 5 adım iken, aynı güzergâh, araya makale noktaları da geldiği için Şekil 1-b’de 10 adım (A-M1-C-M2-D-M3-E-M5-F-M4-G) gibi görünür. Ancak Şekil 1-b’deki, heterojen bilgi ağı, kişilerin çalıştığı bölümleri de içerdiği için homojen sosyal ağda görünmeyen yollar sayesinde A’dan G’ye daha kısa, 6 adımlık (A-B1-D-M3-E-B2-

G) bir yoldan (HBA terminolojisinde meta-yol) ulaşmak mümkün olur. Homojen sosyal ağda da arka planda olan fakat görünmeyen aracı noktalar ve bağlantılar HBA'da görüldüğü için, aslında bu mesafe 6 değil 3 adım olarak düşünülebilir. Yani HBA, kişiler arası bağlantıları daha gerçekçi bir şekilde modellediği için, iletişim maliyeti daha düşük takımlar kurmayı mümkün kılar. Ya da sosyal ağ üzerinde mesafeler çok uzak görüldüğü için ulaşılamayan kişiler nedeniyle tamamlanamayan takımlar heterojen bilgi ağlarında tamamlanabilir. Diğer taraftan, heterojen bilgi ağındaki nokta ve kenarların farklı özelliklerini takım oluşturma yöntemine dahil etmek mümkündür. Örneğin 1-b'deki E ve G kişisi aynı bölümde çalışıyor gibi görünseler de 5 ve 6 numaralı kenar özelliklerine bakarak bu bağlantıların farklı zaman aralıklarına karşılık geldiğini, dolayısıyla bu kişilerin aslında hiçbir zaman birlikte çalışmadığını tespit etmek de mümkün olur. Ya da karar vericinin ortak çalışma için talep ettiği eşik süre kısıtı sağlanmadığı durumda da bu ilişki takım oluşturmada göz ardı edilebilir.

Bu bağlamda çalışmamızın literatüre katkıları şunlardır:

1. Bu çalışmada, gerçek akademik işbirliklerinde etkisi kritik olan aynı kurumda çalışma bağları, birlikte çalışılan süreyi de gözeterek takım oluşturma problemine dahil edilmektedir.
2. Kişilerin uzmanlıkları, ortak yazarı oldukları makalelerin anahtar kelimelerinden alınmaktadır. Ancak burada da kullanıcının belirlediği eşik değerden daha eski çalışmalar (örneğin, üzerinden en az 15 yıl geçmiş makaleler) göz ardı edilmektedir. Dolayısıyla uzmanlığın güncelliği de takım oluşturma problemine dahil edilmektedir.
3. Kişinin aynı anahtar kelimeye sahip makale sayısı o anahtar kelime için yetkinlik düzeyi olarak kabul edilmektedir. Böylece, önerilen sezgisel takım oluşturma yöntemiyle, takım içi iletişim maliyetini enküçüklemeye çalışırken, aynı zamanda takıma seçilen kişilerin yetkinlik düzeyleri toplamını da enbüyüklemeye çalışılmaktadır.
4. Literatürde ilk kez, YÖK Akademik portalından temin edilen verilerle oluşturulan bir heterojen bilgi ağı üzerinde takım oluşturma çalışması gerçekleştirilmektedir.

3. YÖNTEM (METHODS)

Çalışma kapsamında 5 aşamalı bir yöntem takip edilmiştir:

1. aşamada heterojen bilgi ağını oluşturmak üzere uygun veri kaynağı (YÖK Akademik) seçilmiş ve hazırlanan bir web kazıyıcı (web scraper) ile veriler temin edilmiştir.
2. aşamada bu verilerde heterojen bilgi ağını oluşturmak/zenginleştirmek üzere veri düzenleme işlemleri yapılmış ayrıca, OpenAlex [49] kullanılarak verimizde mevcut DOI'si bilinen makaleler için anahtar kelimeler getirilmiştir.
3. aşamada verilerle ağ veritabanı (neo4j) üzerinde HBA oluşturulmuş ve genel ağ istatistikleri (düğüm/kenar sayıları vb.) hesaplanmıştır.
4. aşamada esas amacımıza hizmet eden kurumdaşığa ve zamana duyarlı takım oluşturma algoritması ile birlikte, kıyaslama yapabilmek için, onun daha esnek bir versiyonu ve sadece ortak yazarlığa dayanan bir takım oluşturma algoritması geliştirilmiştir.
5. Önceki aşamaların doğru bir şekilde tamamlandığı teyit edildikten sonra 3 farklı algoritma ile 200 farklı proje (anahtar kelime listesi) için takım oluşturma algoritmaları test edilmiş ve sonuçlar analiz edilmiştir.

3.1. Veri Kaynağı Seçimi ve Veri Temini (Data Source Selection and Data Acquisition)

Literatürdeki gerçekçi ve geniş kapsamlı bilgilere dayanan takımların önemli bir kısmı [9, 26, 40 vb.] bilgisayar bilimi alanındaki bilimsel çalışmaları düzenli bir biçimde barındıran DBLP [50] verisetini kullanmaktadır. Ancak bu verisetinde kişilerin kurum bilgisi mevcut değildir. Bilimsel makalelerde yazarların kurum bilgileri bulunsun da bu veriler de genellikle düzenli bir formatta değildir ve kişinin kurumda hangi yıllar arasında çalıştığına dair net bir bilgi vermez. Çalışmamız kapsamında aynı kurumda çalışma ilişkilerini takım oluşturma probleminde kullanmak istediğimiz için, YÖK Akademik platformunda¹ mevcut verilerin kullanılmasına karar verilmiştir. YÖK Akademik Platformu, Türkiye'deki tüm kamu ve vakıf üniversitelerinde görevli bilim insanlarının hem geriye dönük çalışma bilgilerini hem de yayın bilgilerini içeren ve kamu tarafından idame ettirilen bir arama motoru [51] olduğu için takım oluşturma problemimiz için oldukça uygun bir veri kaynağıdır. Veri temini ve takım oluşturma kodlarının çalışma sürelerini de göz önünde bulundurarak

¹ akademik.yok.gov.tr

platformdaki verinin bir bölümünün kullanılmasına karar verilmiş, hazırlanan kodlarla YÖK Akademik platformundan Endüstri Mühendisliği ve Yönetim Bilişim Sistemleri (YBS) bilim alanlarında görev yapan öğretim elemanlarının profil bilgileri, yayın bilgileri, tez danışmanlık kayıtları, eğitim ve akademik görev geçmiş bilgileri çekilmiştir.

3.2. Veri Ön İşleme (Data Preprocessing)

YÖK Akademik platformundan Endüstri Mühendisliği ve YBS bölümleri için 4 farklı dosya oluşturulmuştur:

master_arastirmaci: Araştırmacıların kimlik bilgileri (ad-soyadı, araştırmacı ID, ORCID, temel alan, bilimsel (alt) alan, uzmanlık alanları)

master_makale: Makale kayıtları (başlık, DOI, yıl, yayın yeri, kapsam, endeks, makale türü, yazar listesi (yazar adı, hex_ID))

master_tez: Araştırmacıların danışmanlık yaptığı doktora/yüksek lisans tez kayıtları (tez türü, yıl, öğrenci, tez başlığı, üniversite)

arastirmaci_gecmis: Akademik geçmiş kayıtları (görev ve öğrenim bilgileri, kurum, birim, unvan, yıl aralıkları)

Tablo 1’de ham veri setlerinin özeti verilmiştir.

*Tablo 1. Ham Veri Setlerinin Özeti
(Summary of Raw Datasets)*

Veri Dosyası	End. Müh.	YBS
master_arastirmaci	1.433	587
master_makale	21.534	7.494
master_tez	4.133	1.034
arastirmaci_gecmis	6.857	3.551

Bir sonraki aşamaya geçmeden önce iki alanın dosyaları birleştirilmiştir.

3.2.1. Yazar Kimlik Eşleştirilmesi (Author Identity Matching)

YÖK Akademik sisteminde makale yazar listeleri, yazarlara atanmış onaltılık (hex) kimlik numaraları ile JSON formatında saklanmaktadır. Ancak bu kimlikler, araştırmacı profil numaralarından (Arastirmaci_ID) farklıdır. İki kimlik sistemi arasında eşleştirme yapabilmek için ilk aşamada, birleşik veri dosyasındaki her makale kaydında ilgili araştırmacının adı ile yazar listesindeki (Yazarlar sütunu) isimler karşılaştırılarak doğrudan eşleştirme yapılmıştır (1.684 eşleşme).

İkinci aşamada, doğrudan eşleşemeyen hex kimlikler için isim ve makale başlığı doğrulaması birlikte

uygulanmıştır: bir hex kimliğinin sahibi ile aynı ada sahip (Arastirmaci_ID’si mevcut) bir araştırmacının, aynı başlıklı en az bir makalede birlikte yer alması durumunda eşleştirme yapılmıştır (11 ek eşleşme).

Son aşama olarak araştırmacıların tez öğrencileri ve ortak yazarları arasında isim benzerliği kontrolü yapılmıştır. Ad soyadı tam eşleşen 16 öğrenci ve ortak yazarlar tekleştirilmiştir.

Bu işlemler sonucunda, toplam 6.464 benzersiz yazardan 1.699’u iç yazar olarak eşleştirilmiş, 4.765’i dış yazar olarak sınıflandırılmıştır. Burada iç yazar YÖK Akademik platformunda, Endüstri Mühendisliği ve Yönetim Bilişim Sistemleri bölümlerinde kayıtlı kişileri, dış yazar ise bu bölümlerde olmayan veya YÖK Akademik platformunda hiç bulunmayan kişileri ifade etmektedir.

3.2.2. Kurumsal Birim Ayırıştırma (Organizational Unit Parsing)

Çalışmamız kapsamında kişilerin aynı kurumda çalışma ilişkilerini de takım oluşturmada dikkate almak istediğimiz için YÖK veritabanından kişilerin ilişkili olduğu kurum bilgileri de çekilmiştir. Akademik geçmiş verisi, görev kayıtları (Tür=Görev) ve öğrenim kayıtları (Tür=Öğrenim) olarak iki türdedir. Görev kayıtları genellikle üniversite, fakülte ve bölüm bilgisi içerirken, eğitim bilgileri çoğunlukla doktora ve lisansüstü mezuniyeti nedeniyle üniversite, enstitü ve program bilgisi içermektedir.

Özellikle büyük üniversitelerde çalışan araştırmacı sayısı çok fazla olacağı için, aynı üniversitede çalışan ancak fakülte ve bölümleri farklı kişiler arasında ilişki potansiyeli oldukça zayıf olacaktır. Bu nedenle, kurumsal bağlantı ilişkilerini mümkün olduğunca hassas bir şekilde kurabilmek amacıyla görev bilgilerindeki kurum adı, alt birimlere ayrıştırılmıştır. Örneğin "GAZİ ÜNİVERSİTESİ MÜHENDİSLİK FAKÜLTESİ ENDÜSTRİ MÜHENDİSLİĞİ BÖLÜMÜ" ifadesi, üniversite= "GAZİ ÜNİVERSİTESİ", fakülte="MÜHENDİSLİK FAKÜLTESİ" ve bölüm="ENDÜSTRİ MÜHENDİSLİĞİ BÖLÜMÜ" olarak ayrıştırılmıştır. Ayrıştırılamayan birimler (meslek yüksekokulları, yabancı kurumlar vb.) üniversite altında tek dal olarak kaydedilmiştir. Öğrenim kayıtları ise yalnızca üniversite düzeyinde ele alınmış ve yalnızca üniversite bilgisi çıkarılmıştır.

3.2.3. Yazarların Uzmanlık Alanlarının Oluşturulması (Formation of Author Skills)

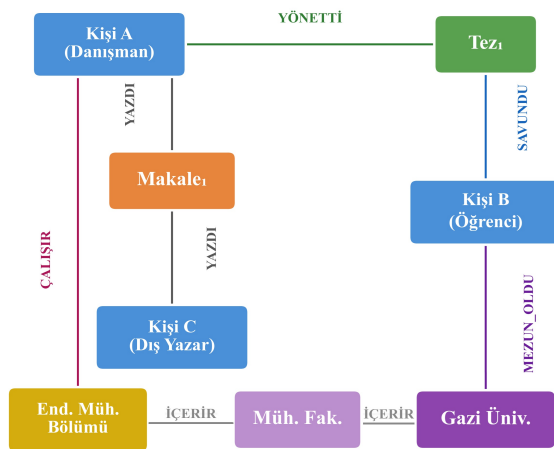
YÖK Akademik platformunda kişilerin "Tedarik Zinciri Yönetimi", "Yöneylem Araştırması" gibi üst-alan düzeyinde, sınırlı (Endüstri Müh. Ve YBS altında toplam 51 çeşit) seçenek arasından uzmanlık alanı

seçmesine müsaade edilmektedir. Bu sınırlılık, kişilerin uzmanlıklarını ayrıştırmaya müsaade etmemektedir. Diğer taraftan platformda, yazarların becerisi olarak kullanılabileceğimiz, makalelerin anahtar kelime bilgisi bulunmamaktadır. Bu eksikliği giderebilmek için, OpenAlex API'si [49] kullanılarak DOI'si bilinen makaleler için anahtar kelimeler getirilmiştir. Toplam 26.374 makalenin 14.795 tanesine OpenAlex API aracılığıyla anahtar kelime bilgisi eklenmiştir. Makale başına gelen anahtar kelime sayısının 11,05 olduğu görülmüştür.

3.3. Ağ Veritabanının Oluşturulması (Formation of Graph Database)

Sosyal ağlar veya heterojen bilgi ağlarından takım oluşturma çalışmalarında sıklıkla kullanılan bir araç Python üzerinde çalışan NetworkX kütüphanesidir. Ancak, veriyi kaydetme, okuma, üzerinde işlemler yapma vb. amaçlarla kullanılabilecek verimli bir araç da ağ veritabanlarıdır. Çalışmamız kapsamında veriyi metin dosyalarında kaydedip NetworkX ile işlemler yapmak yerine, dünyada en çok kullanılan ağ veritabanlarından biri olan Neo4J ile ağ veritabanı oluşturmak tercih edilmiştir. Neo4J ile, seçilen değişkenler düğüm, kenar, düğüm özelliği veya kenar özelliği olarak kaydedilebilmekte, bunların çok çeşitli varyasyonları ile veritabanı işlemleri yapılabilmektedir.

Bir önceki aşamada eşleştirme ve ayrıştırma işlemlerini yaptığımız veri, daha sonra Şekil 2'de görünen düğüm ve kenarları içerecek şekilde ağ veritabanına aktarılmıştır. Aktarım sonrası ağda bulunan düğüm ve kenar sayıları Tablo 2'de verilmiştir. Tezler, tez danışmanlığı ve öğrencilerle ilgili ilişkiler de bu çalışma kapsamında kullanılacak olmakla birlikte ağ veritabanına eklenmiştir.



Şekil 2. Heterojen Bilgi Ağı Şeması
(Heterogeneous Information Network Scheme)

Oluşturulan ağ/graf veritabanında altı düğüm, altı kenar türü bulunmaktadır: Kişi düğümleri araştırmacı kimliği, ORCID, temel_alan, alt_alan, görevler (yazar/danışmanı/tez öğrencisi) özelliklerini; Makale düğümleri DOI, yayın yeri, endeks bilgisi ve OpenAlex API'den elde edilen anahtar kelimeleri (14.795 makale için) taşımaktadır. Tez düğümleri başlık, yıl, tür (yüksek lisans/doktora) ve üniversite bilgisi içermektedir. ÇALIŞIR bağlantısı başlangıç yılı, bitiş yılı, unvan özelliklerini, MEZUN_OLDU bağlantısı mezuniyet derecesi özelliğini içerir.

Tablo 2. Düğüm/Kenar Türleri ve Sayıları
(Node/Edge Types and Counts)

Öge Türü	Tür	Sayı
Düğüm	Kişi	10.649
	Makale	26.374
	Tez	4.989
	Üniversite	637
	Fakülte	1.299
	Bölüm	678
Kenar	YAZDI	51.298
	YÖNETTİ	5.146
	SAVUNDU	5.146
	İÇERİR	1.977
	ÇALIŞIR	4.651
	MEZUN OLDU	5.619

3.4. Takım Oluşturma Algoritmaları (Team Formation Algorithms)

Giriş Bölümünde belirttiğimiz amaçlarımıza uygun olarak, kurum ilişkilerini ve ilişkilerin gerçekleşme zamanını dikkate alan bir sezgisel algoritma tasarlanmış, daha sonra kıyaslama yapılabilmesi için bu algoritmanın 2 varyasyonu daha üretilmiştir.

3.4.1. Aynı Kurumda Çalışma İlişkisinin Zamanını Dikkate Alan Algoritma – ZDA_3_Yıl (Time and Affiliation-Aware Algorithm - ZDA_3_Years)

Algoritmanın en kapsamlı versiyonu olup, şu şekilde çalışmaktadır:

1. Takımda ihtiyaç duyulan yetkinlikler ile kişilerin makalelerindeki anahtar kelimelerin eşleşmesini dikkate alınır. Kişinin bir konuda uzman sayılması için o anahtar kelimeyi içeren en az bir makalede ortak yazar olması gerekir.
2. Kişiler takıma eklenirken verilen anahtar kelime listesindeki sıraya göre uzmanlar aranır. İlk kişi (tohum yazar), ilk anahtar kelime ile makalesi olan yazarlar arasından seçilir.
3. İlgili anahtar kelimeye sahip birden fazla yazar olduğunda, o anahtar kelime ile makale sayısı en fazla olanı seçilir.

4. Uzmanlık bilgisinin güncel olması için, belirlenen bir eşik değer yılından önce (bu çalışmada 15 yıl) yapılmış çalışmalar hesaba katılmaz.

5. İlk yazardan sonra yeni kişiler aranırken öncelikle ortak yazarlık ilişkisi ile arama yapılır. (Yazar_1)-[Yazdı]-(Makale)-[Yazdı]-[Yazar_2] koşulunu sağlayacak bir Yazar_2 bulunamazsa, aynı kurumda çalışma ilişkisi kullanılarak, yani (Yazar_1)-[Çalıştı|İçerir...3]-[Yazar_2] ilişkisiyle en fazla 3 adım uzaklıkta arama yapılır. Burada da biri bulunamazsa (Yazar_1)-[Yazdı|Çalıştı|İçerir...5]-[Yazar_2] dolaylı bağlantıları da dahil edilerek en fazla 5 adım uzaklıkta arama yapılır. Bu son seçenekle bulunan kişiler, çalışmamızda “uzak bağlantı” olarak tanımlanmıştır.

Heterojen Bilgi Ağları literatüründe meta-yol olarak adlandırılan bu tür yollar, sosyal ağ analizine kıyasla mesafeleri daha uzak gibi gösterir. Ancak, (Yazar_1)-[Yazdı]-(Makale)-[Yazdı]-[Yazar_2] yolu bizim ağımızda 2 adım gibi görünse de sosyal ağ analizindeki ortak yazarlık bağlantısına (doğrudan bağlantı) karşılık gelir. Bu çalışma kapsamında izin verdiğimiz toplam 5 adım da sosyal ağ analizinde doğrudan bağlantı yaklaşımı ile 2,5 adıma karşılık gelmektedir.

6. (Yazar_1)-[Çalıştı]-(Bölüm)-[Çalıştı]-(Yazar_2) bağlantısından takıma kişi seçilebilmesi için görev başlangıç ve bitiş tarihleri aralıklarında en az 3 yıl çakışma olması koşulu eklenmiştir. Kullanıcı tarafından değiştirilebilecek bu parametre ile aynı kurumda çalışma ilişkisinin gerçekçiliğinin artırılması sağlanmıştır.

Algoritmanın formel gösterimi Algoritma 1’de sunulmaktadır.

GİRDİ:

$K = \{k_1, k_2, \dots, k_n\}$ ▶ Proje anahtar kelimeleri ($n = |K|$)
 $G = (V, E)$ ▶ Heterojen bilgi ağı (V : noktalar, E : kenarlar)
 T ▶ Yayın güncellik eşiği (yıl)
 $\Delta_{\min}$ ▶ Kurumsal bağlantı zamansal çakışma eşiği (yıl)
 d_{oy}, d_k, d_u ▶ Arama mesafeleri: ortak yazarlık, kurum, uzak

ÇIKTI:

Takım = $\{t_1, t_2, \dots, t_n\}$ ▶ Her a. kelime için bir uzman
 C ▶ Takım içi top. iletişim maliyeti
 U ▶ Takım toplam uzmanlık derecesi

TANIM:

güncel(m) := $(R - \text{yıl}(m)) \leq T$ (m makalesi güncel mi?
 R : referans yılı)
 uzman(a, k_i) := $\exists m : a \in \text{yazarlar}(m) \wedge k_i \in \text{kelimeler}(m) \wedge \text{güncel}(m)$

$\text{çakışma}(a, b) := \min(\text{bitiş}_a, \text{bitiş}_b) - \max(\text{başlangıç}_a, \text{başlangıç}_b)$
 $\text{ms}(a, k_i) := |\{m : a \in \text{yazarlar}(m) \wedge k_i \in \text{kelimeler}(m) \wedge \text{güncel}(m)\}|$
 $[Y|Ç|İ]$: YAZDI|ÇALIŞIR veya İÇERİR kenarı
 (b) : birim (Üniversite/Fakülte/Bölüm) noktası

Algoritma 1:

1. Tohum yazar seçimi

```

2. Adaylar ← {a ∈ V : uzman(a, k1) }
3. t1 ← argmax_{a ∈ Adaylar} ms(a, k1)
4. Takım ← {t1}, C ← 0, U ← ms(t1, k1)
5. FOR i = 2, ..., n DO   ▶ Kalan her anahtar kelime için
6.   üye ← NULL
7.   ▶ Aşama 1: Ortak yazarlık araması
8.   A1 ← {a ∉ Takım : ∃ t ∈ Takım, ∃ m1 : {a, t} ⊆ yazarlar(m1) ∧ güncel(m1) ∧ uzman(a, ki)}
9.   IF A1 ≠ ∅ THEN
10.    üye ← argmax_{a ∈ A1} ms(a, ki)
11.    d ← 2   ▶ d: Kişi-Makale-Kişi
12.    ▶ Aşama 2: Kurumsal bağlantı araması (d ≤ dk)
13.    IF üye = NULL THEN
14.     A2 ← {a ∉ Takım : ∃ t ∈ Takım, yol(t, a) ≤ dk [Ç|İ] ∧ ∃ b : (t)–[Ç]→(b)–[Ç]–(a) ∧ çakışma(t, a, birim) ≥ Δmin ∧ uzman(a, ki)}
15.     IF A2 ≠ ∅ THEN
16.      üye ← argmax_{a ∈ A2} ms(a, ki)
17.      d ← yol_uzunluğu(Takım, üye)
18.    ▶ Aşama 3: Uzak arama (d ≤ du, tüm kenarlar)
19.    IF üye = NULL THEN
20.     A3 ← {a ∉ Takım : ∃ t ∈ Takım, yol(t, a) ≤ du [Y|Ç|İ] ∧ uzman(a, ki)}
21.     IF A3 ≠ ∅ THEN
22.      üye ← a ∈ A3 : min yol(Takım, a), sonra max ms(a, ki)
23.      d ← yol_uzunluğu(Takım, üye)
24.    IF üye ≠ NULL THEN
25.     Takım ← Takım ∪ {üye}
26.     C ← C + d
27.     U ← U + ms(üye, ki)
28.   ELSE
29.     RETURN (Takım, C, U, EKSIK) (Takım eksik)
30. RETURN (Takım, C, U, TAM)

```

3.4.2. Aynı Kurumda Çalışma İlişkisinin Zamanını Dikkate Almayan Algoritma – ZDA Zamansız (Affiliation Time-Unaware Algorithm – ZDA No_Time)

Esas algoritma ile tek farkı, aynı kurumda çalışma ilişkisini değerlendirirken zaman aralığı çakışmasını dikkate almamasıdır. Açık kaynaklarda, kişilerin kurumlarda görev alma bilgileri yer alsa da bunların zaman aralıklarını tespit etmek zordur. Bu algoritma,

çalışılan yıl bilgisinin olmadığı durumlar düşünülerek kurgulanmıştır.

3.4.3. Ortak Yazarlığa Dayalı Algoritma - OYDA (Co-Authorship Dependent Algorithm - OYDA)

Literatürdeki temel takım oluşturma çalışmalarına benzer şekilde, yalnızca ortak yazarlık ilişkilerini göz önünde bulunduran algoritmadır. Bu seçenekte, kurum bağlantıları hesaba katılmaz, yayın yılının çok eski olması da önemsenmez. Yalnızca ortak yazarlık ilişkileri kullanılabilirdi için, takımların tamamlanma ihtimalini artırmak için bu yöntemde izin verilen en uzak mesafe 6 adımdır; yani takımdaki kişinin ortak yazarının ortak yazarı da eklenebilir.

3.5. Takım Oluşturma Testleri (Team Formation Tests)

3 Algoritma varyasyonu aynı anahtar kelime listeleri ve aynı ilk yazarlar ile çalıştırılarak aşağıdaki analizlerin yapılması hedeflenmiştir:

1. Takım Tamamlanma Oranı: Bir takım, proje için tanımlanan k adet anahtar kelimenin her biri için en az bir uzman içerdiğinde "tam takım" olarak kabul edilir. "Tamamlanma oranı", N deney içinde tam takım oluşturulabilen deneylerin oranıdır. Bu tanım doğrultusunda 3 farklı algoritmanın istenen tüm yetkinlikleri barındıran takım oluşturma oranları hesaplanacak ve karşılaştırılacaktır.

2. Takım İçi İletişim Maliyeti: Bu çalışma kapsamında, ihtiyaç duyulan her yetkinlik için takımda mevcut bir elemana en yakın kişi eklenmektedir. Son eklenen iki kişi arasındaki mesafelerin toplanması ile elde edilen takım içi toplam mesafe, iletişim maliyeti olarak tanımlanmıştır. Bu tanım doğrultusunda 3 farklı algoritmanın ürettiği tam takımlarda takım içi iletişim maliyeti hesaplanacak ve karşılaştırılacaktır.

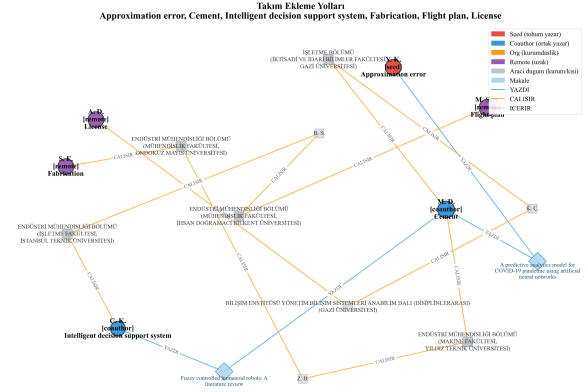
3. Takımların Toplam Yetkinlik Düzeyi: Takımlara seçilen kişilerin ilgili anahtar kelimelerde yaptığı yayınların sayıları toplamı 3 farklı algoritma için hesaplanacak ve karşılaştırılacaktır.

Bu 3 soruya yanıt aramak için öncelikle, her biri ağ veritabanındaki makalelerde mevcut, endüstri mühendisliği ve bilişimle alakalı toplam 4519 anahtar kelimelik bir havuz (Ek-1) oluşturulmuştur. Daha sonra 3 farklı algoritmada testler yapmak üzere, bu havuzdan, farklı projeleri temsilen, her biri 6 kelimeden oluşan 200 farklı anahtar kelime seti oluşturulmuştur. Bir anahtar kelime listesi ile takım oluşturulacak takımın tamamlanma olasılığı o listedeki kelimelerin aday havuzunda bulunma sıklığına da bağlı olduğu için Ek-1 her bir kelimenin ağda (makalelerde) bulunma sıklığını da

göstermektedir. Bu çalışma kapsamında, en az 50 makalede bulunan anahtar kelimelerle takım oluşturma denemeleri yapılmıştır.

4. SONUÇLAR (RESULTS)

Aynı kurumda çalışma ilişkisinin zamanını dikkate alan algoritma ile oluşturulan bir takım, Şekil 3'te görselleştirilmiştir.



Şekil 3. Örnek Takımın Oluşturduğu Alt-Çizge
(Sub-Graph Formed by a Team)

Şekle göre, önce "approximation error" anahtar kelimesi için tohum yazar Y.K. bulunmuş; O'na önce ortak makale ilişkisi (YAZDI) ile M.D., ardından M.D.'nin bir diğer ortak yazarı C.K. eklenmiştir.

Kalan üç anahtar kelime için ortak yazar bulunamadığı için, algoritma kurumsal ve uzak arama aşamalarına geçmiştir: S. E. ve M. S. A.'ya, aynı bölümde çalışan aracı kişileri üzerinden (ÇALIŞIR) 4 adım mesafede ulaşılmıştır. A. D. ise farklı bir fakülte üzerinden yine 4 adımlık bir kurumsal yol ile eklenmiştir. Bu örnek, ZDA algoritmasının ortak yazarlık ağına yetersiz kaldığı durumlarda kurumsal ilişkileri kullanarak takımı tamamlayabildiğini somut olarak göstermektedir.

4.1. Takım Oluşturma Testleri (Team Formation Tests)

6 kelimeden oluşan 200 farklı anahtar kelime seti 3 farklı algoritma kurgusu ile çalıştırılmış ve oluşan takımlara ilişkin bilgiler Ek-2'de anonimleştirilerek paylaşılmıştır. 3 algoritmanın takım oluşturma istatistikleri Tablo 3'teki gibidir.

Tablo 3. 200 Denemelik Test Sonuçları
(Results of 200-Trials Test)

	ZDA Zamansız	ZDA 3 Yıl	OYDA
Deney_Sayısı	200	200	200
Tam Takım Sayısı	189	189	176
Tamamlanma_Oranı	0,945	0,945	0,88

Ort_Maliyet_Tamamlanan	11,77	11,6	12,02
Ort_KW_Makale_Tamaml.	8,19	8,27	8,92
Ort_Takim_Büyükülüğü	5,9	5,9	5,64
Ort_Sure_sn	0,19	0,20	0,50
Ort_Ortak_Yazar_Sayısı	3,7	3,72	4,64
Ort_Org_Bağlantı_Sayısı	0,6	0,4	0
Ort_Uzak_Bağlantı_Sayısı	0,59	0,77	0

Tablo 3'e göre ZDA algoritmasının her iki varyantı da %94,5 tamamlanma oranına ulaşırken, yalnızca ortak yazarlık ilişkisine dayanan yöntemin tamamlanma oranı, daha geniş bir arama mesafesine (6 adım) sahip olmasına rağmen, %88'de kalmıştır. Bu sonuç, kurumsal bağlantıların takım tamamlama kapasitesine anlamlı katkı sağladığını göstermektedir.

Ortalama takım büyüklüğü (bulunan toplam üye sayısı/200) ZDA varyantlarında 5,9 (6 üzerinden) iken, ortak yazarlık yönteminde 5,64'e düşmektedir.

İletişim maliyeti açısından da ZDA varyantları (11,6–11,8) ortak yazarlık yönteminden (12,02) daha düşük ortalama değerlere ulaşmıştır. Bu durum, kurumsal bağlantıların daha kısa/verimli bağlantı yolları sunduğunu ortaya koymaktadır.

Takım oluşturma denemelerinin ortalama süresi sn cinsinden ZDA varyantlarında yaklaşık 0,2 sn; OYDA'da ise 0,5 sn.'dir. Her iki süre de kullanıcı açısından oldukça kısa, kabul edilebilir büyüklükte olsa da, ZDA'nın süre anlamında da daha verimli olduğu anlaşılmaktadır.

İki ZDA varyantı arasındaki fark ise kaynak dağılımında belirginleşmektedir: 3 yıllık çakışma kısıtı uygulandığında kurumsal bağlantı sayısı 0,60'tan 0,40'a düşmekte, buna karşılık uzak bağlantı sayısı 0,59'dan 0,77'ye yükselmektedir. Bu durum, daha sıkı çakışma koşulunun bazı kurumsal ilişkileri eleyerek algoritmayı uzak arama aşamasına yönlendirdiğini, ancak tamamlanma oranını etkilemediğini göstermektedir.

Ortak yazarlık yönteminin daha yüksek ortalama anahtar kelime–makale sayısı (8,92'ye karşı ~8,2) ise bu yöntemin yayın ağı içindeki daha üretken araştırmacıları seçebildiğine, ancak bu dar havuzun takım tamamlama açısından yetersiz kaldığına işaret etmektedir.

Tablo 3'te Tamamlanma_Oranı değişkeninin her iki ZDA versiyonu için aynı değeri vermesi nedeniyle, bu algoritmalar aynı takımları üretiyor olabilir mi sorusu ortaya çıkmıştır. Bu durumu incelemek üzere algoritmaların ürettiği yazar kümelerinin benzerliği toplu olarak ve her bir anahtar kelime listesi bazında incelenmiş, Tablo 4'te tüm test sonuçlarının benzerliği, Tablo 5'te aynı projeye karşılık gelen takımların benzerlik değerleri sunulmuştur.

Tablo 4. Tüm Yazarlar Bazında Benzerlik
(Global Similarity of Author Sets)

Karşılaştırma	Ortak	Birleşim	Jaccard
ZDA-Zamansız - ZDA-3 Yıl	399	419	0.952
ZDA-Zamansız - OYDA	293	479	0.612
ZDA-3 Yıl vs OYDA	287	479	0.599

Tablo 5. Proje Bazlı Takım Benzerliği
(Team Similarity on Project Basis)

Karşılaştırma	Birebir aynı	Ort. Jaccard
ZDA-Zamansız - ZDA-3 Yıl	180/200 (%90)	0.949
ZDA-Zamansız - OYDA	50/200 (%25)	0.530
ZDA-3 Yıl vs OYDA	52/200 (%26)	0.535

Tablo 4 ve 5'teki sonuçları Ek-2'deki takım bilgileri ile birlikte incelediğimizde aşağıdaki sonuçlara ulaşılmaktadır:

- ZDA-Zamansız ve ZDA-3 Yıl'ın önerdiği takımların %90'ı birebir aynıdır. Zamansız çakışma kısıtı çoğu durumda sonucu değiştirmemiştir.
- ZDA versiyonları ve OYDA'nın önerdiği takımların sadece ~%25'i aynı takımdır. Jaccard benzerliğinin de 0,53 olması, yöntemlerin belirgin şekilde farklı yazarlar seçtiğini göstermektedir.
- 119 spesifik yazar yalnızca ZDA yöntemleri ile bulunmakta, OYDA'nın bakmadığı aynı kurumda çalışma bağlantıları aracılığıyla gelmektedir.
- OYDA'nın 67 benzersiz yazarı, mevcut takım üyeleriyle aynı kurumda çalışma ilişkisi bulunmayan, ortak yazarlık zincirinin uzak adımlarından gelen kişilerdir.

5. TARTIŞMA (DISCUSSION)

Proje yöneticisi rolündeki bir bilimsel araştırmacının, verilen uzmanlık alanları ile bir takım oluşturmaya gerektiğinde, öncelikle geçmişte birlikte çalıştığı, dolayısıyla doğrudan bağlantılı olduğu kişilere ulaşması beklenir. Geçmişte birlikte çalıştığı kişilerde bu uzmanlık yoksa, ikinci aşamada bağlantılarının bağlantılarına ulaşması da beklenebilir. Ancak 3. derece bağlantı, yani ortak yazarlarının ortak yazarlarının ortak yazarlarına ulaşmak ve onları proje ekibine dahil etmek nispeten düşük olasılıklı/zor bir süreçtir. Diğer taraftan, kişi geçmişte aynı bölümde/fakültede/üniversitede birlikte çalıştığı kişilerin uzmanlık alanlarına hâkim olabilir ve onlarla tanışıyor olması ihtimali de yüksektir. Bu bakımdan, sosyal ağlar üzerinde takım oluşturma probleminde yalnızca ortak yazarlığın kullanılması mantıksal olarak yetersiz görünmektedir.

Çalışmamızda elde edilen sonuçlar da yalnızca ortak yazarlık ilişkilerine dayalı olarak takım oluşturma, takım tamamlama oranını önemli ölçüde düşürdüğünü göstermektedir. 200 farklı anahtar kelime listesiyle ve ortak yazarlığa 6 adımda arama imkânı verdiğimiz testlerde çıkan %6'lık performans düşüşü yanında benzer bir testi ortak yazarlığa 4 adımda arama imkânı vererek tekrarladığımızda ZDA algoritmalarında başarımlar azalmazken OYDA algoritmasının takım tamamlama oranının %38'e düştüğü görülmüştür. Dolayısıyla deneysel sonuçlarımızın en net bulgusu, kurumsal ilişkilerin takım tamamlama oranını anlamlı düzeyde artırmasıdır. Bu artış, ortak yazarlık bağlantılarının ilgili uzmanlıklara tek başına erişemediği durumlarda kurumsal ilişkilerin (CALISIR, ICERIR) tamamlayıcı bir rol oynamasından kaynaklanmaktadır. göstermektedir. Özellikle niş veya disiplinlerarası uzmanlıklar için, aynı kurumun alt birimlerinde birlikte çalışmış ama daha önce birlikte yayın yapmamış araştırmacıların bulunabilmesi, ZDA algoritmasına önemli bir esneklik kazandırmaktadır.

ZDA'nın kurumsal ilişkilerin çakışma süresini dikkate alan ve almayan iki versiyonunda da ortalama iletişim maliyetinin (11,6–11,8) OYDA yönteminden (12,02) düşük olması da beklenen bir durumdur. Kurumsal hiyerarşi (Bölüm → Fakülte → Üniversite) üzerinden kurulan aynı kurumda çalışma bağlantıları, daha uzun mesafeli ortak yazarlık zincirlerini kontrol etmeye gerek bırakmamakta, doğrudan veya daha kısa erişim yolları sunmaktadır.

ZDA-Zamansız ve ZDA-3 Yıl arasındaki karşılaştırmada, zamansal kısıtların tamamlanma oranına bir etkisi olmadığı ortaya çıkmıştır. ZDA-Zamansız ve ZDA-3 Yıl takımlarının proje bazlı %90'lık benzerliğini de göz önünde bulundurduğumuzda, algoritmaların gerektiğinde farklı kişileri takıma dahil edebildiği, dolayısıyla beklendiği gibi çalıştığı açıktır. Burada tamamlanma oranlarının aynı çıkması ve benzerliğin %90 olması ise algoritmadaki zaman çakışması kısıtının etkisizliğinden değil, YÖK platformundan çekilen verilerin yapısından kaynaklanmaktadır. Yani verimizde, kurumsal bağlarla birbirine bağlanan kişilerin çok büyük bir bölümü en az 3 yıl birlikte aynı kurumda çalışmıştır.

Yine de 3 yıllık çakışma kısıtı uygulandığında takımlara kurumsal bağlantı ile gelen kişi sayısı ortalaması 0,60'tan 0,40'a düşmekte, uzak bağlantı sayısı ise 0,59'dan 0,77'ye yükselmektedir. Farklı bir bakış açısıyla, bu sonuç, algoritmanın daha sıkı çakışma koşulları bazı kurumsal ilişkileri elese de, uzak aramalarla bu eksikliği telafi edebildiğini göstermektedir.

Testlerden ortaya çıkan diğer bir bulgu OYDA yönteminin daha yüksek yetkinlikte (ortalama anahtar kelime-makale sayısı 8,92'ye karşı ~8,2) yazarlar bulabilmesidir. Dolayısıyla, daha uzak mesafelerde arama izni verildiğinde, ortak yazarlık bağlantılarının daha yetkin takım üyeleri bulduğu izlenimi oluşmaktadır. Ancak bu durumun bir sebebi de ZDA yöntemlerinin kişilerin makalelerini kontrol ederken de bir zaman eşiğiyle bakması, bu çalışma özelinde üzerinden 15 yıldan daha fazla geçmiş makalelerdeki anahtar kelimeleri hesaba katmaması olabilir.

KISITLAR (LIMITATIONS)

Bu çalışma YÖK Akademik Platformu'ndan alınan ve sadece Endüstri Mühendisliği ve Yönetim Bilişim Sistemleri ana alanlarında çalışan kişiler temel alınarak hazırlanmıştır. Daha geniş veya farklı bir veri seti ile ortaya çıkacak sonuçlar önemli farklılıklar gösterebilir. Ayrıca kişilerin detaylı uzmanlık bilgileri ve makalelerin anahtar kelimeleri platformda mevcut olmadığı için, bu bilgiler DOI'si mevcut olan makaleler için OpenAlex API'si kullanılarak getirilmiştir. API'nin sağladığı anahtar kelimeler makalelerin ve yazarlarının uzmanlıkları ile tam olarak örtüşmüyor olabilir.

Endüstri Müh. ve YBS alanlarındaki araştırmacıların makalelerindeki diğer kişiler, makalelerin anahtar kelimelerinde uzman olarak çalışmamıza dahil edilmiştir. Ancak bu kişilerin hangi üniversitede/fakültede/bölümde çalıştıkları bilgisi platformda sunulmamıştır. Dolayısıyla heterojen bilgi ağındaki (öğrenciler dahil) 10 binin üzerindeki kişinin yalnızca %16'sının CALISIR ilişkisine sahip olması, aynı kurumda çalışma ilişkilerini de gözetken algoritmanın performansını düşürmüş olması beklenen durumdur. Ağdaki tüm kişilerin çalışma bilgileri temin edilebilirse, ZDA ve OYDA yöntemleri arasındaki performans farkları daha da belirginleşebilir.

Bu çalışma kapsamında, her bir proje (anahtar kelime seti) tanımlanırken en az 50 makalede bulunan anahtar kelimeler havuza dahil edilmiştir; yani çok nadir bulunan uzmanlıklar göz ardı edilmiştir. Daha nadir uzmanlıklar da dahil edilirse takım tamamlanma oranlarının düşmesi beklenir. Ayrıca, her bir proje için 1 takım önerilmiştir. Ancak her proje için en iyi-k takımı bulmak istediğimizde de algoritma performansları değişebilir.

GELECEK ÇALIŞMALAR (FUTURE WORK)

Gelecek çalışmalarda, YÖK Akademik Platformu'ndaki tüm yazar bilgileri alınarak algoritmaların test edilmesi planlanmaktadır. Böylece

ÇALIŞIR ilişkilerinin oranının da artması, ZDA algoritmasının etkisinin daha açık bir şekilde görülmesi beklenmektedir. Ayrıca, bu çalışmada ağ yapısına eklenen ancak takım oluşturmada kullanılmayan lisansüstü tezleri ve bu tezlerle ilgili bağlantıları da göz önünde bulunduran bir yöntemin geliştirilmesi planlanmaktadır.

MADDİ DESTEK (FUNDING)

Yazarlar, çalışma kapsamında herhangi bir kurum, kuruluş veya fon tarafından desteklenmemiştir.

TEŞEKKÜR (ACKNOWLEDGMENTS)

Nuri ÖZDEMİR, doktora tez çalışmalarına imkan verdiği için işvereni TÜBİTAK'a teşekkür eder.

YAZAR KATKI ORANLARI (AUTHORS' CONTRIBUTIONS)

Nuri Özdemir: Kavramsallaştırma, Yöntem Geliştirme, Yazılım, Biçimsel Analiz, Veri Düzenleme, İnceleme, Görselleştirme, Yazım – Orijinal Taslak.

Hadi Gökçen: Danışmanlık, Yazım – Gözden Geçirme ve Düzenleme.

CONFLICT OF INTEREST (ÇIKAR ÇATIŞMASI BEYANI)

Bu çalışmada herhangi bir çıkar çatışması yoktur.

DATA AVAILABILITY (VERİ ERİŞİLEBİLİRLİĞİ)

Çalışmada kullanılan veriler kamuya açık YÖK Akademik Platformu'ndan derlenmiş olup, veriler ek dosya olarak da paylaşılmıştır.

ETHICAL STATEMENT (ETİK BEYAN)

Çalışmada insan veya hayvan katılımcılar yer almamaktadır.

DECLARATION OF AI USAGE (YAPAY ZEKÂ KULLANIM BEYANI)

Çalışma kapsamında YÖK Akademik Platformu'ndan veri çekme amacıyla ve oluşan bir takımı görselleştirmek (bkz. Şekil 3) üzere Python kodlarının hazırlanmasında Claude yapay zekâ aracı kullanılmıştır.

KAYNAKLAR (REFERENCES)

- [1] Stavrou, G., Adamidis, P.A., Papathanasiou, J. B., & Tarabanis, K.A. (2023). Team formation: A systematic literature review[Article]. *International Journal of Business Science and Applied Management*, 18(2), 17–34. doi: 10.69864/ijbsam.18-2.174
- [2] Zakarian, A., & Kusiak, A. (1999). Forming teams: An analytical approach. *IIE Transactions*, 31(1), 85–97. doi: 10.1080/07408179908969808
- [3] Baykasoglu, A., Dereli, T., & Das, S. (2007). Project Team Selection Using Fuzzy Optimization Approach. *Cybernetics and Systems*, 38(2), 155–185. doi: 10.1080/01969720601139041
- [4] Gaston, M.E., Simmons, R., desJardins, M., 2004. Adapting Network Structure for Efficient Team Formation. Technical Report. AAAI Technical Report WS-04-02. Menlo Park, CA. In AAAI Technical Report (WS-04-02), pp. 1-8.
- [5] Fitzpatrick, E.L., Askin, R.G., 2005. Forming effective worker teams with multi-functional skill requirements. *Computers & Industrial Engineering* 48, 593–608. doi:10.1016/j.cie.2004.12.014.
- [6] Backstrom, L., Huttenlocher, D., Kleinberg, J., Lan, X., 2006. Group formation in large social networks: Membership, growth, and evolution, in: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Philadelphia PA USA. pp. 44–54. doi:10.1145/1150402.1150412.
- [7] Cheatham, M., Cleereman, K., 2006. Application of Social Network Analysis to Collaborative Team Formation, in: International Symposium on Collaborative Technologies and Systems (CTS'06), IEEE, Las Vegas, NV, USA. pp. 306–311. doi:10.1109/CTS.2006.18.
- [8] Wi, H., Oh, S., Mun, J., Jung, M., 2009. A team formation model based on knowledge and collaboration. *Expert Systems with Applications* 36, 9121–9134. doi:10.1016/j.eswa.2008.12.031.
- [9] Lappas, T., Liu, K., Terzi, E., 2009. Finding a team of experts in social networks, in: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Paris France. pp. 467–476. doi:10.1145/1557019.1557074.

- [10] Li, C.T., Shan, M.K., 2010. Team Formation for Generalized Tasks in Expertise Social Networks, in: 2010 IEEE Second International Conference on Social Computing, IEEE, Minneapolis, MN, USA. pp. 9–16. doi:10.1109/SocialCom.2010.12.
- [11] Dorn, C., Dustdar, S., 2010. Composing Near-Optimal Expert Teams: A Trade-Off between Skills and Connectivity, in: Meersman, R., Dillon, T., Herrero, P. (Eds.), On the Move to Meaningful Internet Systems: OTM 2010. Springer Berlin Heidelberg, Berlin, Heidelberg. volume 6426, pp. 472–489. doi:10.1007/978-3-642-16934-2_35.
- [12] Kargar, M., An, A., 2011. Discovering top-k teams of experts with/without a leader in social networks, in: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, ACM, Glasgow Scotland, UK. pp. 985–994. doi:10.1145/2063576.2063718.
- [13] Anagnostopoulos, A., Becchetti, L., Castillo, C., Gionis, A., & Leonardi, S. (2010). Power in unity: Forming teams in large-scale community systems. In *Proceedings of the 19th ACM international conference on Information and knowledge management* (pp. 599–608). Toronto ON Canada: ACM. doi: 10.1145/1871437.1871515
- [14] Anagnostopoulos, A., Becchetti, L., Castillo, C., Gionis, A., & Leonardi, S. (2012). Online team formation in social networks. In *Proceedings of the 21st international conference on World Wide Web* (pp. 839–848). Lyon France: ACM. doi: 10.1145/2187836.2187950
- [15] Majumder, A., Datta, S., Naidu, K., 2012. Capacitated team formation problem on social networks, in: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Beijing China. pp. 1005–1013. doi:10.1145/2339530.2339690.
- [16] Rangapuram, S.S., Bühler, T., Hein, M., 2013. Towards realistic team formation in social networks based on densest subgraphs. In: Proceedings of the 22nd International Conference on World Wide Web. ACM, Rio de Janeiro, Brazil, pp. 1077–1088. doi:10.1145/2488388.2488482.
- [17] Awal, G.K., Bharadwaj, K.K., 2014. Team formation in social networks based on collective intelligence – an evolutionary approach. *Applied Intelligence*, 41, 627–648. doi:10.1007/s10489-014-0528-y.
- [18] Mahajan, Y., Cho, J.H., Chen, I., 2024. Privacy-preserving and diversity-aware trust-based team formation in online social networks. *ACM Transactions on Intelligent Systems and Technology*, 15. doi:10.1145/3670411.
- [19] Selvarajah, K., Zadeh, P.M., Kobti, Z., Palanichamy, Y., Kargar, M., 2021. A unified framework for effective team formation in social networks. *Expert Systems with Applications*, 177, 114886. doi:10.1016/j.eswa.2021.114886.
- [20] Yin, X., Qu, C., Wang, Q., Wu, F., Liu, B., Chen, F., Chen, X., Fang, D., 2018. Social connection aware team formation for participatory tasks. *IEEE Access*, 6, 20309–20319. doi:10.1109/ACCESS.2018.2819992.
- [21] Gutiérrez, J.H., Astudillo, C.A., Ballesteros-Pérez, P., Mora-Melià, D., Candia-Véjar, A., 2016. The multiple team formation problem using sociometry. *Computers & Operations Research*, 75, 150–162. doi:10.1016/j.cor.2016.05.012.
- [22] Daş, G.S., Altınkaynak, B., Göçken, T., Türker, A.K., 2022. A set partitioning based goal programming model for the team formation problem. *International Transactions in Operational Research*, 29, 301–322. doi:10.1111/itor.13022.
- [23] Campêlo, M.B., Figueiredo, T.F., 2021. Integer programming approaches to the multiple team formation problem. *Computers & Operations Research*, 133, 105354. doi:10.1016/j.cor.2021.105354.
- [24] Fathian, M., Saei-Shahi, M., Makui, A., 2017. A new optimization model for reliable team formation problem considering experts' collaboration network. *IEEE Transactions on Engineering Management*, 64, 586–593. doi:10.1109/TEM.2017.2715825.
- [25] Cao, N., Lin, Y.R., Li, L., Tong, H., 2015. G-Miner: Interactive visual group mining on multivariate graphs. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, Seoul, South Korea, pp. 279–288. doi:10.1145/2702123.2702446.
- [26] Hu, C., Zhou, Q., Tong, H., 2024. Genius: Subteam replacement with clustering-based graph neural networks. In: Proceedings of the 2024 SIAM International Conference on Data Mining (SDM), pp. 10–18. doi:10.1137/1.9781611978032.2.
- [27] Li, L., Tong, H., Cao, N., Ehrlich, K., Lin, Y.R., Buchler, N., 2016. TEAMOPT: Interactive team optimization in big networks. In: Proceedings of the 25th ACM International Conference on Information and Knowledge Management. ACM, Indianapolis, USA, pp. 2485–2487. doi:10.1145/2983323.2983340.

- [28] Zhou, Q., Li, L., Tong, H., 2019. Towards real-time team optimization. In: 2019 IEEE International Conference on Big Data. IEEE, Los Angeles, USA, pp. 1008–1017. doi:10.1109/BigData47090.2019.9006078.
- [29] El-Ashmawi, W.H., Ali, A.F., Tawhid, M.A., 2019. An improved particle swarm optimization with a new swap operator for team formation problem. *Journal of Industrial Engineering International*, 15, 53–71. doi:10.1007/s40092-018-0282-6.
- [30] El-Ashmawi, W.H., Ali, A.F., Słowik, A., 2020. An improved Jaya algorithm with a modified swap operator for solving team formation problem. *Soft Computing*, 24, 16627–16641. doi:10.1007/s00500-020-04965-x.
- [31] El-Ashmawi, W.H., Sheta, A.F., Alqadi, B.S., Abdelminaam, D.S., Mohammed Alsekait, D.M., 2025. A new version of the golden eagle optimizer algorithm and its application for solving a trio-objective skillful team formation problem in a social network. *Journal of Artificial Intelligence and Soft Computing Research*, 15, 357–384. doi:10.2478/jaiscr-2025-0018.
- [32] Rehman, M.Z., Zamli, K.Z., Almutairi, M.S., Chiroma, H., Aamir, M., Kader, M.A., Mohd Nawi, N., 2021. A novel state space reduction algorithm for team formation in social networks. *PLOS ONE*, 16. doi:10.1371/journal.pone.0259786.
- [33] Sun, Y., Han, J., 2013. Mining heterogeneous information networks: A structural analysis approach. *ACM SIGKDD Explorations Newsletter*, 14, 20–28. doi:10.1145/2481244.2481248.
- [34] Cai, D., Shao, Z., He, X., Yan, X., Han, J., 2005. Community mining from multi-relational networks. In: *European Conference on Principles of Data Mining and Knowledge Discovery*. Springer, pp. 445–452.
- [35] Shafie, T., 2015. A multigraph approach to social network analysis. *Journal of Social Structure*, 16, 1–21. doi:10.21307/joss-2019-011.
- [36] Yang, J., Li, M., Wu, B., Xu, C., 2016. Forming a research team of experts in expert-skill co-occurrence network of research news. In: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 1143–1150. doi:10.1109/ASONAM.2016.7752383.
- [37] Yang, X., Wang, P., Fang, W., 2018. Team formation with relationship strength based on meta-path in heterogeneous network. In: 2018 IEEE 9th International Conference on Software Engineering and Service Science. IEEE, pp. 797–801. doi:10.1109/ICSESS.2018.8663775.
- [38] Hamidi Rad, R., Fani, H., Kargar, M., Szlichta, J., & Bagheri, E. (2020, October). Learning to form skill-based teams of experts. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 2049-2052). doi:10.1145/3340531.3412140.
- [39] Hamidi Rad, R., Bagheri, E., Kargar, M., Srivastava, D., Szlichta, J., 2021. Retrieving skill-based teams from collaboration networks. In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2015–2019. doi:10.1145/3404835.3463105.
- [40] Hamidi Rad, R., Nguyen, H., Al-Obeidat, F., Bagheri, E., Kargar, M., Srivastava, D., Szlichta, J., Zarrinkalam, F., 2023. Learning heterogeneous subgraph representations for team discovery. *Information Retrieval Journal*, 26, 8. doi:10.1007/s10791-023-09421-6.
- [41] Lee, O.J., Hong, S., Kim, J.T., 2021. Interinstitutional research team formation based on bibliographic network embedding. *Mobile Information Systems*, 2021. doi:10.1155/2021/6629520.
- [42] Sharaf, M.A., Alghamdi, T.G., 2025. Balanced team formation using hybrid graph convolution networks and MILP. *Applied Sciences*, 15, 2049. doi:10.3390/app15042049.
- [43] Dara, M., Rad, R.H., Zarrinkalam, F., Bagheri, E., 2025. Retrieval-augmented neural team formation. In: *Advances in Information Retrieval*. Springer Nature Switzerland, pp. 362–371. doi:10.1007/978-3-031-88714-7_35.
- [44] Liu, X., Wu, K., Liu, B., Qian, R., 2023. HNERec: Scientific collaborator recommendation model based on heterogeneous network embedding. *Information Processing & Management*, 60, 103253. doi:10.1016/j.ipm.2022.103253.
- [45] Gao, X., Wu, S., Xia, D., Xiong, H., 2023. Topic-sensitive expert finding based solely on heterogeneous academic networks. *Expert Systems with Applications*, 213, 119241. doi:10.1016/j.eswa.2022.119241.
- [46] Guan, K., Huang, W., Chen, T., Chan, W.K.V., 2025. TEDRec: Transformer-based scientific collaborator recommendation via textual-edge dynamic network modeling. *Information Processing*

& *Management*, 62, 104283.
doi:10.1016/j.ipm.2025.104283.

[47] Wang, X., Huang, W., Zhao, B., Li, S., 2026. Scientific collaborator recommendation via hypergraph embedding. *Information Processing & Management*, 63, 104423.
doi:10.1016/j.ipm.2025.104423.

[48] Liu, J., Zhang, Q., Xu, S., Huang, C., Gong, X., 2026. Innovation performance improvement path in scientific research teams: A perspective of network embeddedness. *Information Processing & Management*, 63, 104484.
doi:10.1016/j.ipm.2025.104484.

[49] Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.

[50] Ley, M. (2002). The DBLP Computer Science Bibliography: Evolution, Research Issues, Perspectives. In: Laender, A.H.F., Oliveira, A.L. (eds) String Processing and Information Retrieval. SPIRE 2002. Lecture Notes in Computer Science, vol 2476. Springer, Berlin, Heidelberg.
https://doi.org/10.1007/3-540-45735-6_1

[51] Anadolu Ajansı, “YÖK Akademik arama motoru kullanıma açılıyor,” 3 Şubat 2016. [Çevrimiçi].
<https://www.aa.com.tr/tr/egitim/yokakademik-arama-motoru-kullanima-aciliyor/530885>