



GAU
GIRNE AMERICAN
UNIVERSITY

Volume 7 - Issue 11

2015

JOURNAL

SOCIAL AND APPLIED SCIENCES

TRNC

Girne American University Journal of Social and Applied Sciences

Owner : Girne American University

Editor : Asst.Prof.Dr. İbrahim Erşan

Advisory Board : Prof.Dr. Sadık Ülker
Assoc.Prof.Dr. Zafer Ağdelen

Cover Graphic Design : Asst.Prof.Dr. İbrahim Erşan

Review Board : Prof.Dr. Sadık Ülker
Prof.Dr. Ali Zeki
Assoc.Prof.Dr. Zafer Ağdelen
Assoc.Prof.Dr. Ali Haydar
Assoc.Prof.Dr. Kamil Dimililer
Assoc.Prof.Dr. Bülent Bilgehan
Assoc.Prof.Dr. Kabir Sadeghi
Asst.Prof.Dr. İbrahim Erşan
Asst.Prof.Dr. Burcu Toker
Asst.Prof.Dr. Tamer Tulgar
Asst.Prof.Dr. Ezgi Deniz Ülker
Asst.Prof.Dr. Mehmet Okaygün
Asst.Prof.Dr. Sara Kandulu
Asst.Prof.Dr. Meryem Erbilek
Asst.Prof.Dr. Necip Suat Canarlan
Dr. Ömer Damdelen
Dr. Zalihe Yarkıner
Dr. Eser Gemikonaklı
Buğçe Eminağa
Sabriye Topal

Mailing Address : GAU Journal of Social and Applied Sciences
Karmi Kampus
Karaoğlanoğlu,
Girne, North Cyprus Via
Mersin 10, Turkey

E-Mail : gaujournal@gau.edu.tr

Web : http://www.gau.edu.tr/online/sosyal_gau_dergisi_ve_uygulamali_bilimler

The ideas published in the journal belong to the respective authors.
Dergide yayımlanan yazıların sorumluluğu yazarlarına aittir.

CONTENTS / İÇİNDEKİLER

| | |
|---|-----|
| Doppler Radarları için Dalgalanma Kayıplarının Algılanan Radar Gücüne Etkisi M. Öcalan, N. Akçam | 1 |
| Fuzzy logic based pesticide sprayer for smart agricultural drone A. Baba | 11 |
| Smart template matching algorithm using hill-climbing search Strategy A. Baba | 19 |
| Analyzing the effect of the slit depth and width on the electrical performance of a squirrel cage induction motor A. G. Yetgin, M. Turan, E. Unlukaya | 26 |
| Underwater Visual Tracking and Counting of Fishes Md.H. Sharif, Ş. Uyaver, A. Güler | 38 |
| Estimation of Targets from Satellite Images A. Youssef, Ş Uyaver, Md.H. Sharif | 45 |
| A Meta-Heuristic Approach for Optimal Train Control K. Keskin, A. Karamancioglu | 51 |
| A Hybrid Algorithm for Exam Timetabling Problem in Marmara University M. E. Gedikli, M. Agaoglu | 57 |
| RSRP: Risk Sensitive Routing Protocol in Wireless Sensor Networks M. Karakaya | 68 |
| Using Service Oriented Architecture for Plate Recognition by Mobile Devices M. Karakaya, G. Şengül | 76 |
| Determining Firm Logos Using Image Processing Techniques on a Mobile Device to Aid Blind People M. Karakaya, G. Şengül | 82 |
| Efficient Data Gathering in WSN with a Range Constrained Mobile Relay I. Cereci, H. Dağlayan, N. Kılınç, S. Aktaş, M. Karakaya | 89 |
| Differential Search Algorithm with Levy Flight E. Sertel, O. Altun | 96 |
| Evaluation of Semantic Similarity Measurement Algorithms For Word Sense Disambiguation Ş. Altun, E. Domnori | 112 |
| Energy Optimization of the Current-Limiting Power LED Drivers in Various Powers Ö. F. Farsakoğlu, H. Y. Hasırcı, İ. Atik, İ. Çelik | 122 |
| Determination Power Correction Parameters of Buck and Buck Boost LED drivers for LED lighting Ö. F. Farsakoğlu, İ. Çelik, İ. Atik, H. Y. Hasırcı | 133 |
| Performance Simulation of Gossip Relay Protocol in Multi-hop Wireless Networks A. Sarı, E. Çağlar | 145 |
| Using Data Mining on Linked Data Y. Gültepe | 152 |
| Kidney Segmentation From Abdominal CT Images By Using Connected Component Labeling Algorithm S. A. Tuncer, A. Alkan | 157 |
| Artificial Neural Network Based Power Distribution System Modelling and Harmonic Estimation S. Özdemir, M. Demirtaş, S. Aydın | 163 |
| Comparison with different models of bending stress analysis of the cantilever beams under different profile section, materials and loads J. Gattmah, M. T. Özkan, İ. Toktaş, E. Demir | 173 |
| Comparison with different models of tensile and compressive stress analysis on a cantilever beam model E. Demir, İ. Toktaş, M. T. Özkan, J. Gattmah | 183 |
| A Novel Method for Islanding Detection of Distributed Generation Units C. F. Okwose , R. Sirjani | 195 |

Doppler Radarları için Dalgalanma Kayıplarının Algılanan Radar Gücüne Etkisi

M.Öcalan¹, N.Akçam²

^{1,2}Elektrik-Elektronik Mühendisliği Bölümü, Gazi Üniversitesi

Eti Mh., Yükseliş Sk., Maltepe, Ankara, TURKEY

¹m_ocalan@yahoo.com, ²ynursel@gazi.edu.tr

ÖZET

Bu makalede, bilgisayar yazılımı MATLAB aracılığıyla bir hedef, bir faz dizili antene sahip monostatik darbeli Doppler radarı ve bu radarın hedefe gönderdiği sinyalin parametreleri girilerek, hedeften yansıyan sinyale ait çıktıların modellenmesi çalışılmıştır. Bu modelleme esnasında kullanılan radar ve hedef parametreleri, var olan birer örneğe uygun olarak alınmış olup, hedefin hareket modellenmesi ve bu modellemenin gerçeğe uygunluğu için ayrı bir modelleme çalışması yapılmıştır. Ayrıca Swerling durumları; radar kayıpları, gürültünün olmadığı ve gürültünün mevcut olduğu durumlar için, modelleme çıktıları karşılaştırılarak etkileri gözlemlenmiştir.

Anahtar Kelimeler – Doppler radarı, Swerling durumları

1. GİRİŞ

Radarı temel amacı düşman ve dost objeler hakkında bilgi toplamak iken, gizliliği yok eden bu teknolojinin karşısında ise karıştırıcı teknolojisi geliştirilmiştir. Radar teknolojisi gönderilen sinyallerin hedeften yansıyıp algılanmasıyla hedefi görmeyi amaçlarken, karıştırıcı teknolojisi sinyalleri bozarak kendisini radar sistemlerine göstermemeyi amaçlar. Bu teknolojilerdeki sinyallerin bir hedeften yansımaya en büyük etkenlerden biri, o hedefin elektromanyetik dalga olan radar sinyallerini yansıtabilirliği olarak tanımlanabilen Radar Kesit Alanı (RKA)'dır. Hedeften yansıyan sinyallerin gücü RKA ile orantılıdır. Radardan gizlenmek için RKA azaltma çalışmaları yapan Sıla Beyhan [1], çeşitli cisimlerin RKA analizlerini yaparak RKA azaltma metodlarını detaylı olarak incelemiştir. Çalışmada hem monostatik hem de bistatik RKA ile birlikte mükemmel elektrik iletkenleri ve radar emici madde ile kaplanılan cisimlerin RKA'ları karşılaştırılmıştır. Bu sayede hedef radar sistemi sinyalleri yansıtılmasında daha küçük gibi davranacak ve yansıyan sinyallerin gücü, radar algılama gücünün altında kalacaktır. Ancak radar teknolojisinde, hedef düşük güç yansıtırsa bile, çeşitli yöntemlerle (gürültünün ve kaybın indirgenmesi veya daha fazla güce sahip sinyaller gönderilmesi gibi) hedefin algılanması sağlanmaya çalışılır. Farine, Lombardo ve Ortenz [2], çalışmalarında yüksek ve alçak kazançlı ışın karışımlarıyla genel dizi anten konfigürasyonları bulunan modern radar sistemlerinin hedef algılama ve yön tayini konularını incelemektedirler. Zeynep Arslan [3] ise, Weibull ve Gauss dağılımlı ortamlarda, N darbeli radar alıcılarının Swerling-I hedef tipi için başarımlarını incelemiştir. Başarım kıyaslamaları; Weibull ortamda, ortama ait şekil parametresinin farklı değerleri için incelenmiş ve Gauss ortam başarımı ile kıyaslanmıştır.

2. RADAR MENZİL DENKLEMİ

Sinyalin antenden hedefe giderken aldığı yol (gidiş menzili) R_g ve hedeften antene ulaşırken aldığı yol ise (dönüş menzili) R_a olmak üzere, etkin anten alanını da içeren bistatik radar denklemi Eş. 1 ve Eş. 2'de verilmiştir. Bu değerlerin aynı olduğu durumlarda ortak bir menzil olan R ifadesi kullanıldığında, monostatik radar denklemi (Eş. 3) elde edilir.

$$P_a = \frac{P_g G \sigma}{4\pi R_g^2} AK_a \quad (1)$$

$$P_a = \frac{P_g G \sigma}{(4\pi)^2 R_a^2 R_g^2} AK_a \quad (2)$$

$R_g = R_a = R$ ise (monostatik radar):

$$P_a = \frac{P_g G \sigma}{(4\pi)^2 R^4} AK_a \quad (3)$$

Burada;

P_a : Hedefte yansıtıp alınan güç (W),

P_g : Gönderilen güç (W),

σ : Radar kesit alanı (m^2),

G : Anten kazancı,

A : Anten yüzey alanı (m^2),

R : Anten-hedef arası uzaklık (m),

K_a : Anten yansıtma katsayısı'dır [4].

2.1 Dalgalanma kayıpları (Swerling durumları)

Radar sistemlerinde hedef, iki durum dikkate alınarak modellenir.

- Radar kesit alanı değişim hızı,
- Genlik değişimlerinin olasılık yoğunluk işlevidir [3].

Radar kesit alanı; hedefin şekline göre yansımının hedefin hangi bölgesinden gerçekleştiğine, hedefin hareketine ve radarın görüş açısına bağlıdır.

Swerling modeli, hedef salınımlarını (genlik değişimleri) tanımlamada yaygın olarak kullanılan bir metot olarak bilinir. Swerling modelinde hedefler salınım hızlarına ve şekillerine göre sınıflara ayrılmıştır. Swerling hedeflere ait radar kesitlerindeki dalgalanma durumları dört farklı olasılık yoğunluk fonksiyonuyla tanımlanır [5]. Bu fonksiyonlar;

Swerling durumu (case) 1, yavaş salınımlı hedefler için tanımlanan durumdur. Swerling-I durumunda hedeflerin genlik dağılımı Rayleigh bağıntısıyla tanımlanır. Bu durum için olasılık yoğunluk fonksiyonu $p(\sigma)$ Eş. 4'de verilmiştir.

$$P(\sigma) = \frac{1}{\sigma_{ort}} e^{-\frac{\sigma}{\sigma_{ort}}} \quad (4)$$

Burada; σ_{ort} , hedefin ortalama RKA değeridir.

Swerling durum (case) 2, Swerling-I durumundan tek farkı, hedefin salınımlarının hızlı olmasıdır. Swerling-II durumunda olasılık yoğunluk fonksiyonu olarak Swerling-I durumundaki fonksiyon kullanılır (Eş. 4).

Swerling durum (case) 3, yavaş salınımlı hedefler için tanımlanan durumdur. Swerling-III durumunda hedeflerin genlik dağılımı, Rician bağıntısıyla tanımlanır. Olasılık yoğunluk fonksiyonu ise Eş. 5 ile ifade edilir.

$$P(\sigma) = \frac{4\sigma}{\sigma_{ort}^2} e^{-\frac{2\sigma}{\sigma_{ort}}} \quad (5)$$

Swerling durum (case) 4, Swerling-IV durumunda hedeflerin genlik dağılımı, Rician bağıntısıyla tanımlanır. Swerling-III durumundan tek farkı hedefin salınımlarının hızlı olmasıdır.

Swerling durum (case) 5 (Swering-0), herhangi bir salınımlı olmayan, RKA değeri sabit olan hedefler için tanımlanan durumdur [5].

3. HEDEF VE RADAR MODELLEMELERİ

Bu çalışmada hedef ve radar sistemi için bilgisayar yazılımında (MATLAB) birer modelleme oluşturulmuş ve gerçeğe yakın şartlar uygulanmaya çalışılmıştır. Bu modellemelerde amaç, sabit bir radar sistemi ile bu radar sistemi kapsam alanından geçen rastgele hareketli bir hedefi canlandırmaktır. Daha sonra radar sistemi ile hedefe sinyal gönderip, yansıyan sinyalin güç değerleri; sistem kayıpları yok iken ve sistem kayıpları var iken incelenmiştir. Ayrıca kayıplı ve kayıpsız durumlar Swerling durumları etkileri altında da incelenmiştir.

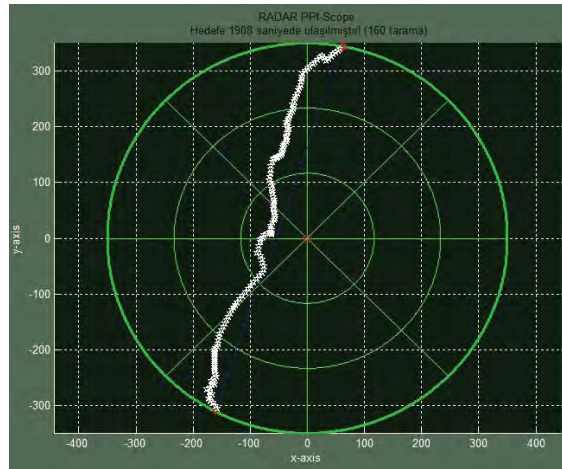
3.1 Hedef Modellemesi, Girdileri ve Çıktıları

Hedef modellemesinde rastgele bir güzergah seçilerek, fakat belirli kurallar çerçevesinde ve kurallara uygun bir hareket algoritmasıyla birlikte bir F-16 Fighting Falcon savaş uçağının bir noktadan başka bir noktaya ulaşırken radar menzili içindeki hareketleri canlandırılmaya çalışılmıştır. Hedefin RKA, radara olan uzaklığı (konum) ve hızı (Doppler kayması), radar modellemesini etkileyecek hedef parametreleridir. Uçağın modellenmesinde kullanılacak gerçek değerlerine yakın olarak seçilen bu parametrelere (Tablo 1) ek olarak menzili yaklaşık 350 km olarak alınmış ve radar sisteminin PPI ekranı görünümü Şekil 1'de verilmiştir.

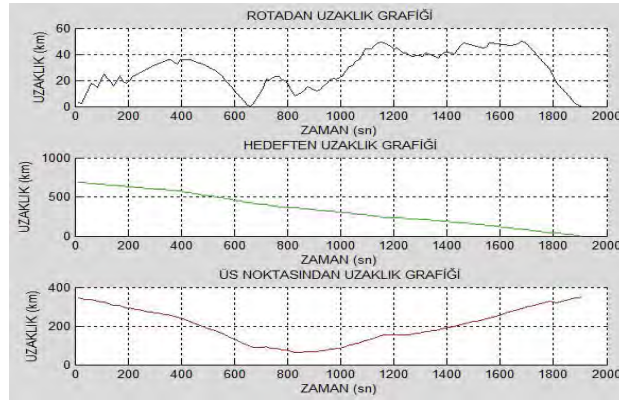
Tablo 1. Hedef modellemesinde kullanılan parametreler

| Parametre adı | Parametre değeri |
|----------------------|----------------------|
| Radar kesit alanı | 5,012 m ² |
| Başlangıç yüksekliği | 10 km |
| Azami yükseklik | 12 km |
| Azami hız | 2400 km/saat |
| Asgari hız | 100 km/saat |
| Azami ivme | 9 g |

Şekil 1'de görüldüğü gibi hedefin radar menziline girişi ve çıkışı arasında 1908 saniye geçmiştir ve bu sürede 160 adet radar taraması gerçekleşmiştir. Modellenen hedefin radar kapsamı içinde olacağı bu süreçte; en kısa rotadan uzaklık, varış noktasından uzaklık ve radardan uzaklık grafikleri Şekil 2'de verilmektedir.

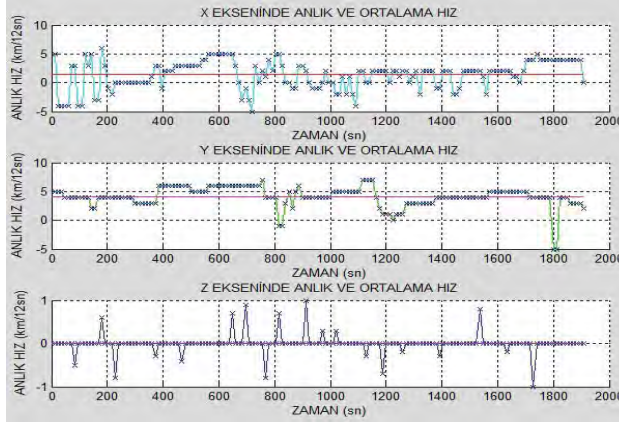


Şekil 1. Hedef modellemesi rota çıktısı



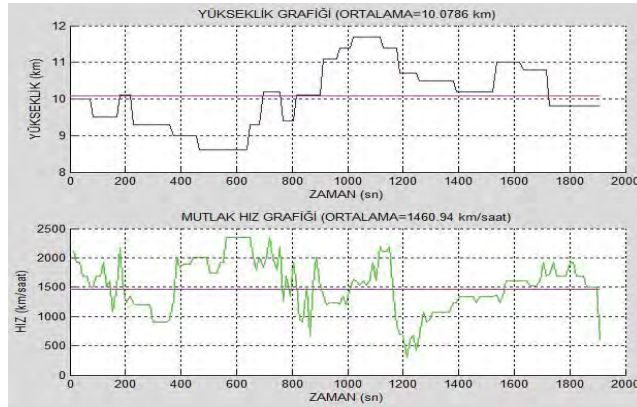
Şekil 2. Hedef uzaklıkları

Modellenen hedefin varış noktasına doğru yol alırken elde edilen x-y-z eksenlerine göre anlık hız ve ortalama hız grafikleri Şekil 3'de görüldüğü gibidir.



Şekil 3. Hedefin eksenlere göre anlık ve ortalama hızları

Modellemede yer alan uçağın radar kapsam alanı içinde bulunduğu sırada sahip olduğu yükseklik ve mutlak hız grafikleri Şekil 4'de gösterilmiştir. Buna göre; hedefin ortalama yüksekliği 10,0786 km ve ortalama mutlak hızı 1460,94 km/saat'tir.



Şekil 4. Hedef yüksekliği ve mutlak hızı

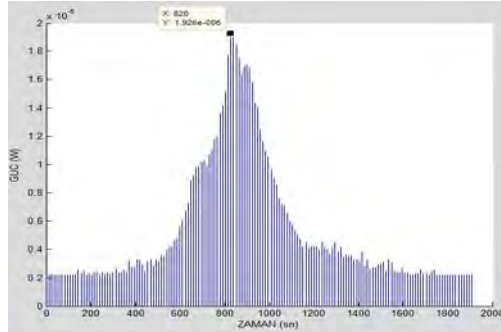
3.2 Radar Sistemi Modellemesi, Girdileri ve Çıktıları

Radar sistemi modellemesinde parametreleriyle birlikte faz dizi antene sahip bir HADR radarı (Hughes Air Defense Radar - Hughes Hava Savunma Radarı) canlandırılmaya çalışılmıştır. Hedef modellemesinden elde edilen gerekli çıktılardan yanı sıra, radar sistemi modellemesinde kullanılacak olan gerçek değerlerine yakın seçilen parametreler Tablo 2'de verilmiştir.

Tablo 2. Radar sistemi modellemesinde kullanılan parametreler [4].

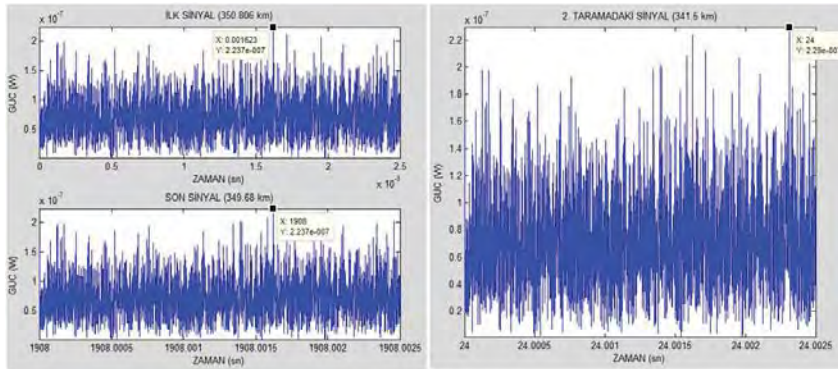
| Parametre adı | Parametre değeri |
|-------------------------------------|------------------|
| Radar çalışma frekansı | 3 GHz |
| Darbe gücü | 2 GW |
| Darbe tekrarlama frekansı | 400 Hz |
| Darbe genişliği | 60 μ s |
| Yaklaşık azami menzil | 350 km |
| Anten tam dönüş süresi | 12 sn |
| Anten kazancı | 30 dB |
| Anten çalışma frekansı | 1-10 GHz |
| İsabet adedi | 1 |
| Algılama olasılığı | 0,9 |
| Yanlış alarm olasılığı | 10^{-6} |
| Gürültü azami gücü (algılama eşiği) | 0,2237 μ W |
| Çevre sıcaklığı | 20 °C |

3.2.1 Modellemede Gürültü Olmaması Durumu. Ortalama güç, darbe tekrarlama zamanı, alım süresi, ölü süre ve kör menzil değerleri gibi parametreler ise yazılımda hesaplanmaktadır [4]. Ayrıca, radar sistemi modellemesinde teoride genellikle kullanılmayan Doppler kayması (kör hız) ve yeryüzü eğikliği de gerçeğe uygunluk açısından göz önünde bulundurulmuştur. Bu değerler ve hedef modellemesinden ortaya çıkan değerlerin girdi olarak kullanılması ve Eş. 3 kullanılarak elde edilen radar sistemi modellemesinin tüm sinyalleri içeren genel çıktısı Şekil 5'te verilmiştir.



Şekil 5. Radar sistemi modellemesinde elde edilen sinyal çıktıları

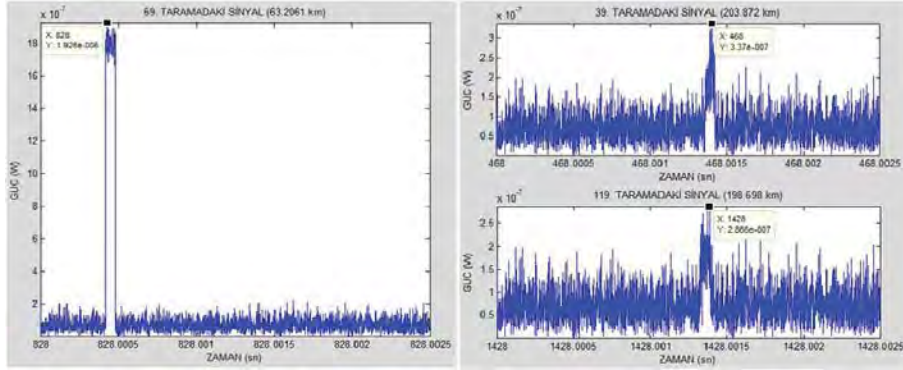
Şekil 6'da 350 km uzaklıktaki hedefe gönderilen ilk ve son sinyallerin yanı sıra, hedef 341,5 km uzaklıkta iken gürültüden kaynaklanan algılama eşiği üzerinde elde edilen algılanan ilk sinyalin çıktıları verilmektedir.



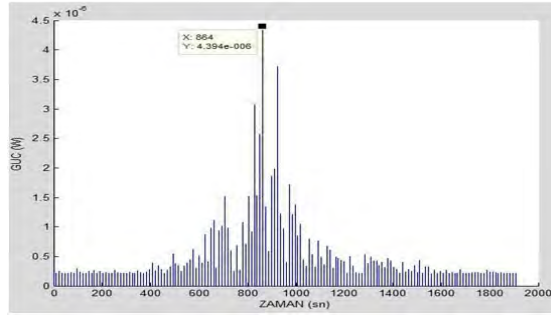
Şekil 6. Hedefin 350 km mesafedeki sinyaller ile algılanabilen ilk sinyal çıktıları

Şekil 6'da, radar sistemi modellemesinde kullanılan gürültü sinyalinin azami değeri 0,2237 μ W'dır ve bu değer aynı zamanda hedeften yansıyan bir sinyalin algılanabilmesi için sahip olacağı asgari

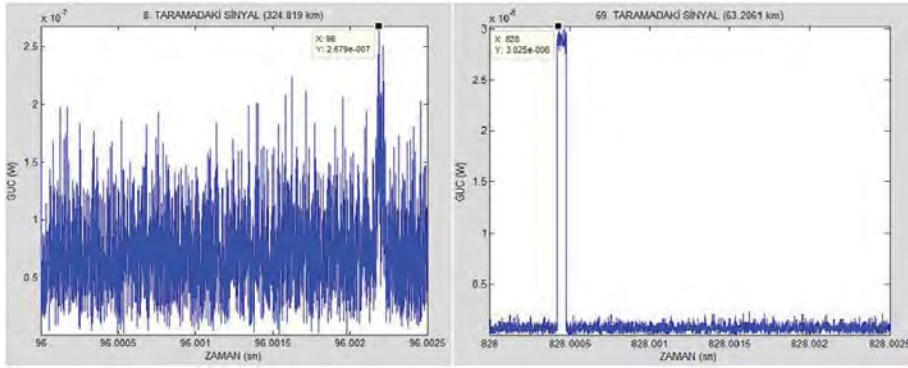
değerdir. Bu nedenle, bu değeri algılama eşiği olarak adlandırmak mümkündür. Şekil 6'da görüldüğü gibi hedef 341,5 km uzaklıktayken algılama eşiği üzerinde algılanabilen ilk değer $0,229\mu W$ olarak elde edilmiştir. Şekil 7'de ise hedef RKA'nda herhangi bir dalgalanma öngörülmediğinden dolayı, hedef radar sistemine en yakın noktadayken (69. tarama, 828. saniye, 63,2061 km) elde edilen en güçlü sinyal ($1,926\mu W$) ve rastgele seçilen 39. ($203,872\text{ km}$) ve 119. ($198,698\text{ km}$) radar taramasındaki sinyal çıktıları görülmektedir.



Şekil 7. Hedef en yakın noktadayken algılanan sinyal ve rastgele sinyal çıktıları

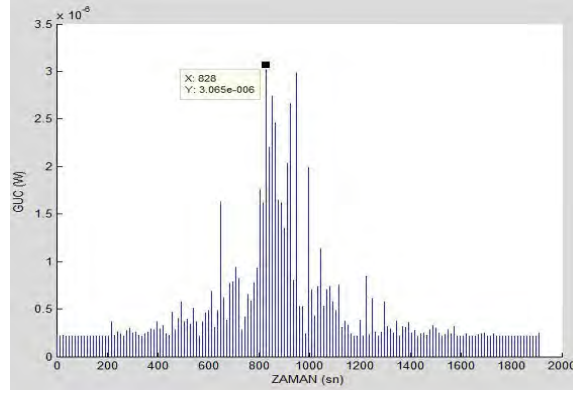


Şekil 8. Swerling-I durumunda elde edilen sinyal çıktıları



Şekil 9. Swerling-I durumunda algılanabilen ilk sinyal ve en yakın noktada algılanan sinyal çıktıları

Modellenen radar sisteminin yaklaşık menzili daha önce belirtildiği gibi 350 km olmakla beraber bu menzil değeri; gürültü, kayıplar ve Swerling durumlarının ortaya çıkması halinde değişmesi beklenen bir değerdir [6-8]. Modellemenin bu aşamasında radar sistemi modellemesine aynı girdiler ile daha önce belirtilen Eş. 4 ve Eş. 5 yardımıyla Swerling durumları uygulanarak yazılımdan algılanan sinyallerin çıktıları elde edilmiştir. Swerling-I durumunda Eş. 4 yardımıyla yavaş salınımlı hedefe göre elde edilen çıktılar Şekil 8'de verilmiştir Şekilde algılanan en yüksek güç değerinin her zaman hedef radara en yakın noktadayken elde edilmediği gösterilmiştir. Ayrıca Swerling-I durumunda elde edilen en güçlü sinyal 864. saniyede $4,394\mu W$ gücü ile 66 km'de elde edilmiştir. Şekil 9'da Swerling-I durumunda algılama eşiğini geçen ilk sinyal ($0,269\mu W$) ve en yakın noktadaki sinyalin ($3,025\mu W$) çıktıları verilmiştir. Buna göre radarın menzilin Swerling-I durumu göz önüne alındığında $324,819\text{ km}$ 'ye düştüğü görülmüştür.

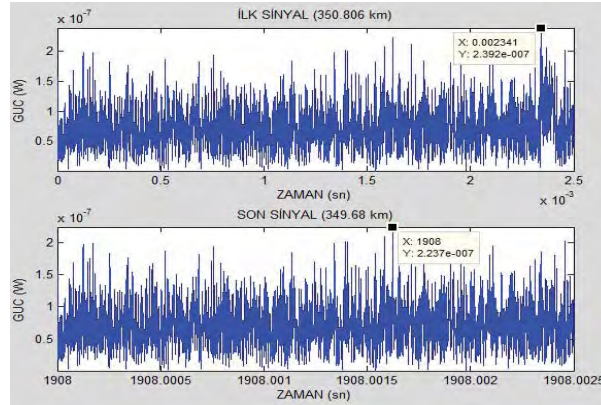


Şekil 10. Swerling-II durumunda elde edilen sinyal çıktıları

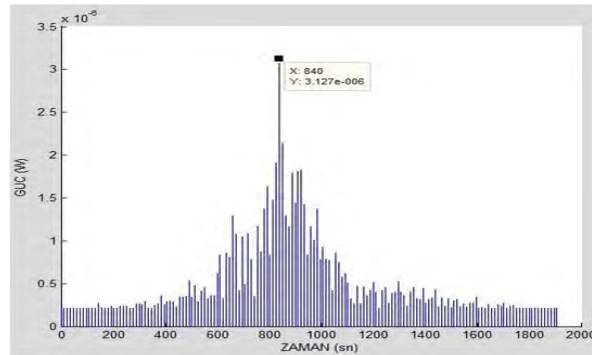
Eş. 4 yardımıyla elde edilen hızlı salınım (hedefin RKA göre) Şekil 10 ile gösterilmiştir. Şekil 10'da Swerling-II durumunun, Swerling-I durumuna kıyasla daha sık salınımlı olduğu ve en yüksek güç değeri olasılıklar dahilinde hedefin radara en yakın olduğu noktada $3,065 \mu\text{W}$ olarak elde edildiği görülmüştür.

Şekil 11 Swerling-II durumunda hedefin 350 km mesafedeyken algılanabildiğini gösteren ilk sinyal çıktısını vermektedir. Bu mesafedeyken algılanan sinyalin yansımaları, $0,2392 \mu\text{W}$ ile algılama eşik değerini aşmış ve radarın hedefi en uzak noktada algılamasını sağlamıştır.

Swerling-III durumunda salınımlar Swerling-I durumundaki gibi yavaştır (Şekil 12), ve faydalanılacak olan bağıntı Eş. 5'tir. Yavaş fakat Swerling-I durumuna göre daha güçlü salınımların gözlemlendiği bu durumda en güçlü sinyal, $3,127 \mu\text{W}$ olarak 63,6 km'de elde edilmiştir.

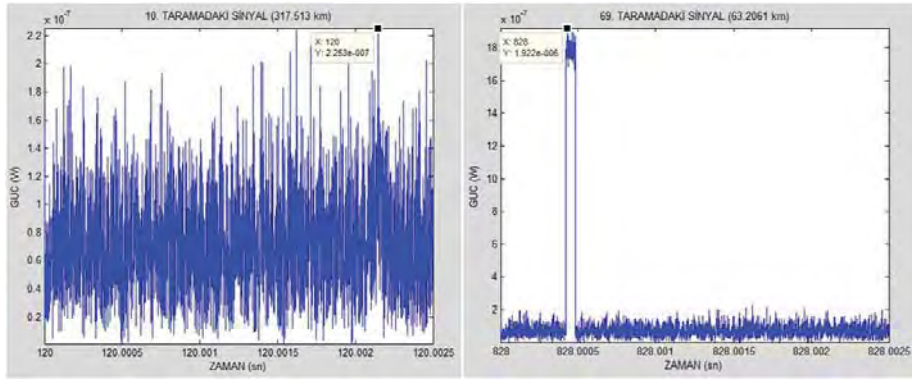


Şekil 11. Swerling-II durumunda 350 km mesafede elde edilen sinyal çıktıları

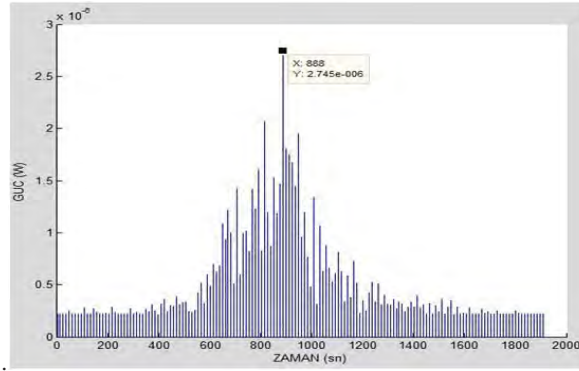


Şekil 12. Swerling-III durumunda elde edilen sinyal çıktıları

Şekil 13'te Swerling-III durumunda algılama eşikini geçen ilk sinyal ($0,2253 \mu\text{W}$) ve en yakın noktadaki sinyalin ($1,922 \mu\text{W}$) çıktıları verilmiştir. Buna göre radarın menzilin, Swerling-I durumu göz önüne alındığında 317,513 km'ye düştüğü görülmüştür

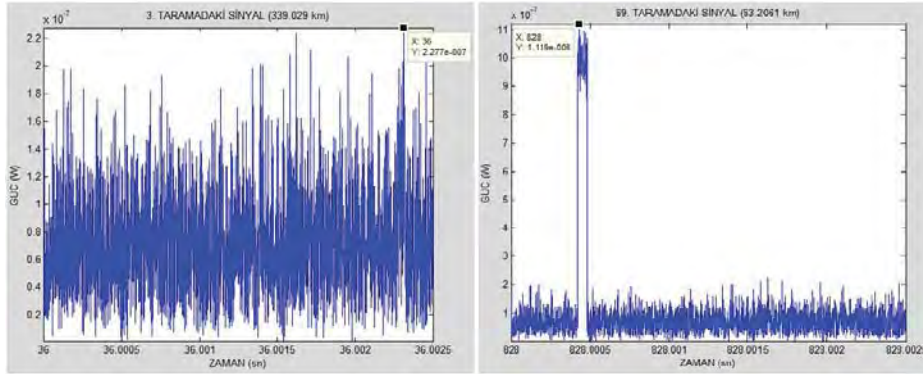


Şekil 13. Swerling-III durumunda algılanabilen ilk sinyal ve en yakın noktada algılanan sinyal çıktıları



Şekil 14. Swerling-IV durumunda elde edilen sinyal çıktıları

Swerling-IV durumu modellemesinde (Şekil 14), Swerling-III durumunda olduğu gibi Eş. 5'den yararlanılmıştır, fakat RKA sınımları Swerling-III durumuna kıyasla daha hızlıdır. Şekil 14'e göre Swerling-IV durumunda algılanan en yüksek güç 2,745 μ W olarak 67,875 km'de elde edilmiştir. Ayrıca algılama eşiğini ilk aşan güç 0,2277 μ W ve en yakın noktada algılanan güç 1,119 μ W'dır (Şekil 15). Swerling-IV durumunda algılama eşiğini ilk aşan güç 339,029 km'de elde edilmiştir. Bu değer radar menzili olarak düşünülebilir.



Şekil 15. Swerling-IV durumunda algılanabilen ilk sinyal ve en yakın noktada algılanan sinyal çıktıları

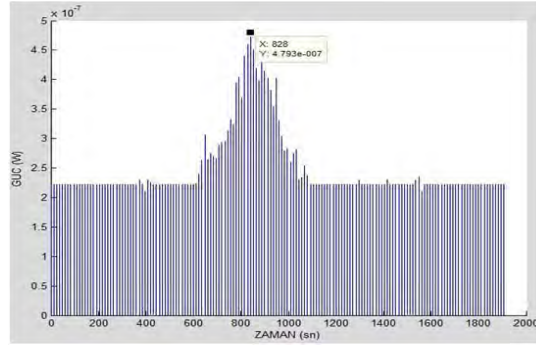
3.2.1 Modellemede Gürültü Olması Durumu. Radar sistemi modellemesinde önceki bölümde, dalgalanma kayıpları (Swerling durumları) göz önünde bulundurulmuş, ancak radar sistemi kayıplarının (atmosferik, sinyal biçiminden kaynaklanan, bant genişliği, filtre uyumlama, toplama, sinyal işleme, alıcı, verici ve anten kayıplarının) sıfıra yakın olduğu varsayılmıştır. Fakat gerçekte durum bundan farklıdır. Bu nedenle radar sistemi modellemesinde Tablo 3'de verilen değerlere göre tipik toplam kayıp değeri (Eş.6) dalgalanma kayıpları dahil yaklaşık 15 dB olarak yazılıma uygulanmıştır. Toplam sistem kaybı,

$$L_{top} = L_{atm} + L_{ss} + L_{BW} + L_m + L_f + L_i + L_{sp} + L_{rx} + L_{tx} + L_{ant} \quad (6)$$

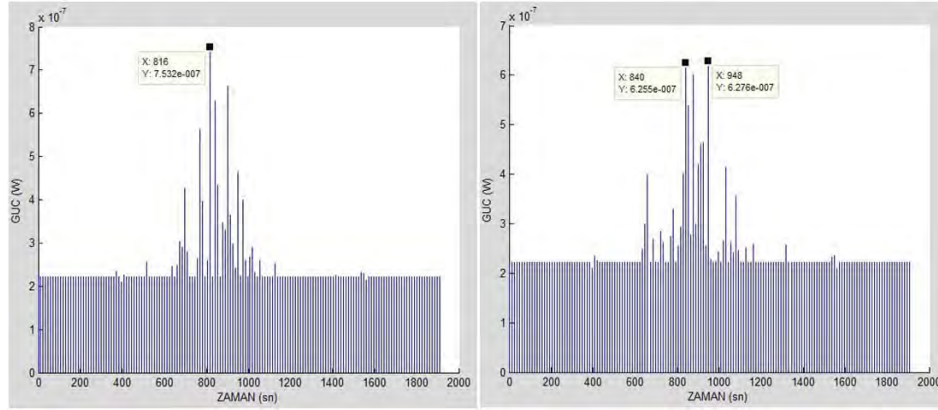
olmak üzere hesaplanır. Burada, L_{ant} : Antenden kaynaklanan kayıplar (dB)'dir.

Tablo 3. Radar sistemi tipik kayıp değerleri

| Kayıp türü | Sembol | Tipik değer |
|--|-----------|-------------|
| Atmosferik kayıplar | L_{atm} | 1,2 dB |
| Sinyal biçiminden kaynaklanan kayıplar | L_{ss} | 1,3 dB |
| Bant genişliği kayıpları | L_{BW} | 1,2 dB |
| Filtre uyumlama (matching) kayıpları | L_m | 0,8 dB |
| Dalgalanma kayıpları ($P_d=0,9$ için) | L_f | 8,4 dB |
| Toparlama (integration) kayıpları | L_i | 3,2 dB |
| Muhtelif sinyal işleme kayıpları | L_{sp} | 3,0 dB |
| İletim hattı kayıpları (alıcı) | L_{rx} | 1,0 dB |
| İletim hattı kayıpları (verici) | L_{tx} | 1,0 dB |
| Tipik toplam sistem kaybı | L_{top} | 21,1 dB |



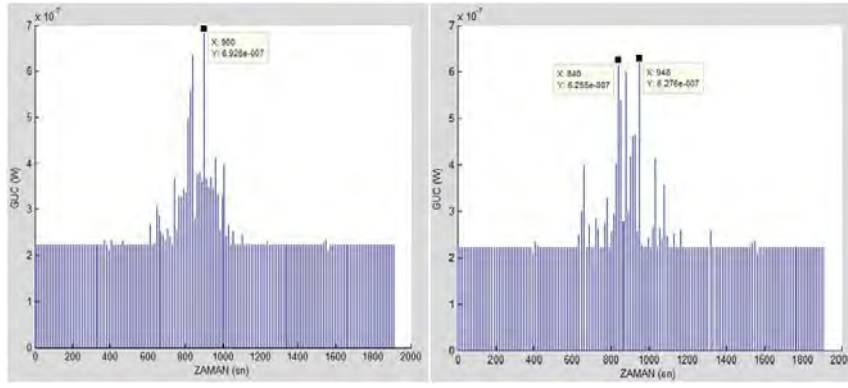
Şekil 16. Radar sistemi modellemesinde kayıplar göz önüne alındığında elde edilen sinyal çıktıları



Şekil 17. Radar sistemi modellemesinde Swerling-I ve Swerling-II durumlarında kayıplar göz önüne alındığında elde edilen sinyal çıktıları

Radar sistemi modellemesinin dalgalanma (Swerling) durumları hariç, kayıplı sinyal çıktıları Şekil 16 ile verilmiştir. Şekle göre, sistemde algılama eşliğinin yükselmemesine rağmen, algılanan gücün kayıplardan kaynaklı olarak oldukça düşmesinden (en yakın noktada $0,4793 \mu W$) dolayı radar menziline bir hayli azaldığı görülmektedir. Bu durumda atmosferik koşullar başta olmak üzere kayıpların radar menziline etkisi Şekil 16'da açıkça görülebilmektedir.

Kayıplarla birlikte Swerling-I ve Swerling-II durumları uygulandığında elde edilen sinyal çıktıları Şekil 17, Swerling-III ve Swerling-IV durumları uygulandığında elde edilen sinyal çıktıları Şekil 18'de görülmektedir. Şekillerde elde edilen sinyal çıktılarındaki dalgalanmalar, olasılıklara bağlı olduğundan önceden verilen kayıpsız modellemelerle benzerdir, fakat aynı değildir.



Şekil 18. Radar sistemi modellemesinde Swerling-III ve Swerling-IV durumlarında kayıplar gözüne alındığında elde edilen sinyal çıktıları

4. SONUÇLAR

Bu çalışmada; Swerling durumları, kayıplar, gürültü ve Doppler kayması gibi durumların radar menziline etkileri bilgisayar yazılımı vasıtasıyla canlandırılmaya çalışılarak bir radar sistemi modellemesi geliştirilmiştir. Ayrıca modelleme için gerekli olan hedef bilgilerinin girilmesi için gerçeğe uygun bir hedef modellemesi geliştirilmiş, çıktıları da radar sistemi modellemesi hesaplamalarında kullanılmıştır. Bu iki modellemede de mevcut bir radar sistemi ve bir uçağın teknik karakteristiklerinden yararlanılmıştır. Radar sistemi modellemesinde, radar sistemi girdileri ile hedef modellemesinden gelen hedef girdileri hesaplanarak radarın algıladığı hedeften yansıyan güç sinyallerinin çıktıları elde edilmiştir. Daha sonra radar sistemine Swerling durumları uygulanmış ve bu durumlarda algılanan güçler önceden elde edilen güçlerle karşılaştırılarak radar menzili üzerindeki etkileri incelenmiştir. Daha sonra radar sistemi üzerine kayıplar uygulanmış ve çıktıları diğer sinyal çıktıları ile karşılaştırılarak radar menzili üzerindeki etkileri gözlemlenmiştir. Son olarak da Swerling durumları ve kayıplar aynı anda radar sistemi modellemesine uygulanarak elde edilen çıktılar analiz edilmiştir. Ayrıca her aşamada radar sistemi modellemesinde gürültünün ve bundan kaynaklı oluşan asgari algılama gücünün de radar menzili üzerindeki etkileri gözlemlenmiştir. Modellemeler sonucu, gürültünün doğrudan algılama eşiğini etkilemesinden dolayı radar menzili sınırladığı görülmüştür. Aynı şekilde kayıpların da algılanan gücü azalttığından radar azami menzili azalttığı gözlemlenmiştir. Swerling durumlarında ise salınımlar ve bu salınımların hızları algılanan güçlerde açıkça görülmüştür.

5. KAYNAKLAR

- [1] A. Farina, P. Lombardo, L. Ortenzi, "A unified approach to adaptive radar processing with general antenna array configuration", Signal Processing, 2004, pp. 84: 1593-1623.
- [2] S. Beyhan, "Feko simülasyon programı ile rastgele şekilli nesnelerin radar kesit alanlarının simülasyonu", Lisans Tezi, Kocaeli Üniversitesi Mühendislik Fakültesi, Kocaeli, 2011, pp. 1-2, 22-23.
- [3] Z. Arslan, "Çok darbeli radar sinyallerinin Gauss ve Weibull ortamlarda sezimi", Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2006, pp. 1-6, 95-97.
- [4] M. Öcalan, "Doppler Radarları için Dalgalanma Kayıplarının Algılanan Radar Gücüne Etkisi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2013.
- [5] M. I. Skolnik, "Introduction to Radar Systems 2nd ed.", Mc Graw Hill, Singapore, 1981, pp. 1-3, 7-8, 15-98.
- [6] H. Meikle, "Modern Radar Systems 2nd ed.", Artech House, Norwood, 2008, pp. 447-488.
- [7] R. B. Mahafza, "Radar Systems Analysis and Design Using Matlab", Chapman & Hall / CRC, Alabama, 2000, pp.118-120.
- [8] D. K. Barton, "Modern Radar Systems Analysis", Artech House, Norwood, 1988, pp. 60-85.

Fuzzy logic based pesticide sprayer for smart agricultural drone

Abdellatif BABA

Mechatronics department, University of Turkish Aeronautical Association,
Bahçekapı Quarter Okul Street, Etimesgut, ANKARA, Turkey

ababa@thk.edu.tr

ABSTRACT

Agricultural drones are becoming modern tools to provide farmers with a lot of details about their crops and to achieve some precise farming missions at specific times for special types of plants and for limited areas. Pesticide spraying is one of the most important chemical agricultural applications. A new generation of smart unmanned aerial vehicles is discussed in our paper to achieve an efficient and economic spraying operation.

In this paper we present our study for a fuzzy logic based sprayer, its functionality depends on several parameters like: the drone altitude, the drone speed, the wind speed and the green surface density which is evaluated using some techniques of digital image processing. The drone is supposed to be equipped by a camera, inertial unit, FPGA card, GPS module, integrated wireless card, and an ultrasound sensor for determining the drone altitude.

Keywords – Fuzzy logic controller (FLC), pesticide spraying, unmanned aerial vehicle (UAV), chemical agricultural application, green surface density evaluation.

1. INTRODUCTION

Thanks to its facility to be built without the need for an accurate mathematical model and to its flexibility to be modified by the on-site staff. A fuzzy logic controller is proposed in this paper to adjust the quantity of pesticides which have to be applied over an agricultural crop by an unmanned aerial vehicle. Several factors represent the inputs of our controller (the drone altitude and speed, the wind speed and finally the green surface density which has to be determined using some techniques of digital image processing).

You et al. [1] studied the extent of the crop disease stress and acoustic emission, the relationship between environmental factors based on the transpiration rate temperature, humidity, light intensity and CO₂ concentration as a crop input precision spraying system. Miller et al. [2] presented an experiment to determine the effectiveness of using a UAV for dispersing pesticides to reduce human disease due to insects. The purpose of that study was to perform non chemical or least toxic chemical techniques to control pests and disease vectors. Huang, et al. [5] have designed an agricultural spray system for small UAV including specialized electrostatic rotary atomizers. The use of multiple coordinated UAVs for spray application was presented by Wang, et al. [7]. Yan et al. [4] proposed a study of map based automatic spraying pesticide system to targets in real time sensory technology. The performance of their system was not sufficiently strong. The requirement for low volume application, in consideration of limited payload capacity, was proposed by Ru et al. [6]. Sugiura et al. [3] achieved the development of on-board monitoring systems to assist the ground based observer about the situation of the UAV's status.

The objective of our study is to develop a spraying system for an autonomous UAV that can precisely apply sprays for agricultural products protection purposes. Detailed explanations about the fuzzy logic controller will be presented in the next paragraph. Then, evaluating the green surface density from each given colored image will be discussed in the third paragraph. The construction of the sprayer will be described with some practical considerations in the fourth paragraph. Finally, we conclude with some perspectives.

2. FUZZY LOGIC

Depending on some rules that describe how a system is working, the behavior of its output will be changed for each difference in its input. This is in general the structure of any control system. These last rules are normally represented by mathematical models which become more difficult to be formulated when the complexity of the system increases. While in fuzzy logic, the mathematical model is replaced by fuzzy rules which are written using some linguistic terms (fuzzy model).

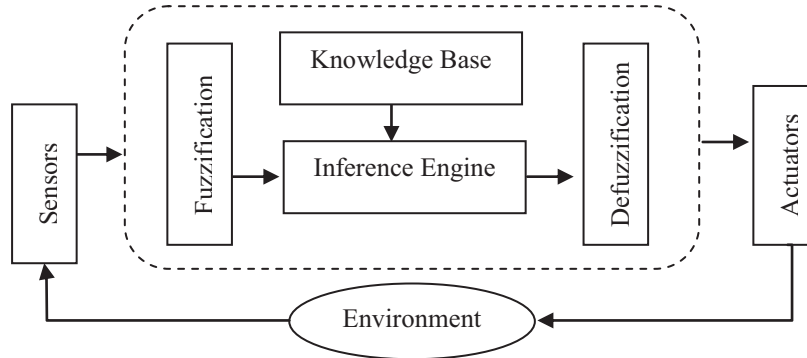


Figure 1. The general structure of any fuzzy logic controller

2.1 Fuzzification step, membership functions

In this first step, any given inputs have to be converted from crisp logic values to fuzzy logic values (fuzzification) as illustrated in the Fig. 1. Then, the inference engine that contains the control rules (conditional rules are written using some pre-defined and saved terms in the data base) will determine the correct output for each group of those inputs. The determined fuzzy output has to take its value in crisp logic (defuzzification) to be finally applied on the actuators. The functionality of the converting operations (fuzzification and defuzzification) depends on some pre-selected membership functions that are playing the main role of giving each crisp input its fuzzy meaning in the fuzzification step and the vice versa in the defuzzification step.

In this paper, the membership functions illustrated in Fig. 2 are adopted to achieve the fuzzification. In each function four main fuzzy sets are distinguished (Low, Medium, High and Very High). The vertical axe in all the successive functions represents an evaluation function as a normalized value between zero and one. The horizontal axe in each one represents a different crisp input. For our specifications, the drone altitude will be considered (Low) for those heights between zero and fifty meters, (Low Medium) between fifty and one hundred meters, (Medium) between one hundred and two hundred meters, (Medium High) between two hundred and three hundred meters, (High) between three hundred and five hundred meters, (More High) between five hundred and one thousand meters and finally it will be considered (Very High) for those heights which are more than one thousand meters. For the other membership functions, the same principle will be employed. The drone speed limits are considered between zero and more than four hundred twenty kilometer per hour. The wind speed limits are considered between zero and more than ninety kilometer per hour. And, the percentage of green surface density will be extracted from the green component of the successive captured colored images that are describing the current scenes.

2.2 Decision making

The last explained step was responsible of giving each crisp input its own meaning in linguistic terms (Fuzzy sets). Then, a new step of decision making has to be done using the inference engine which is normally composed of successive conditional rules (if and then expressions). For some designing purposes, arranging all those expressions in some tables will be clearer and more practical.

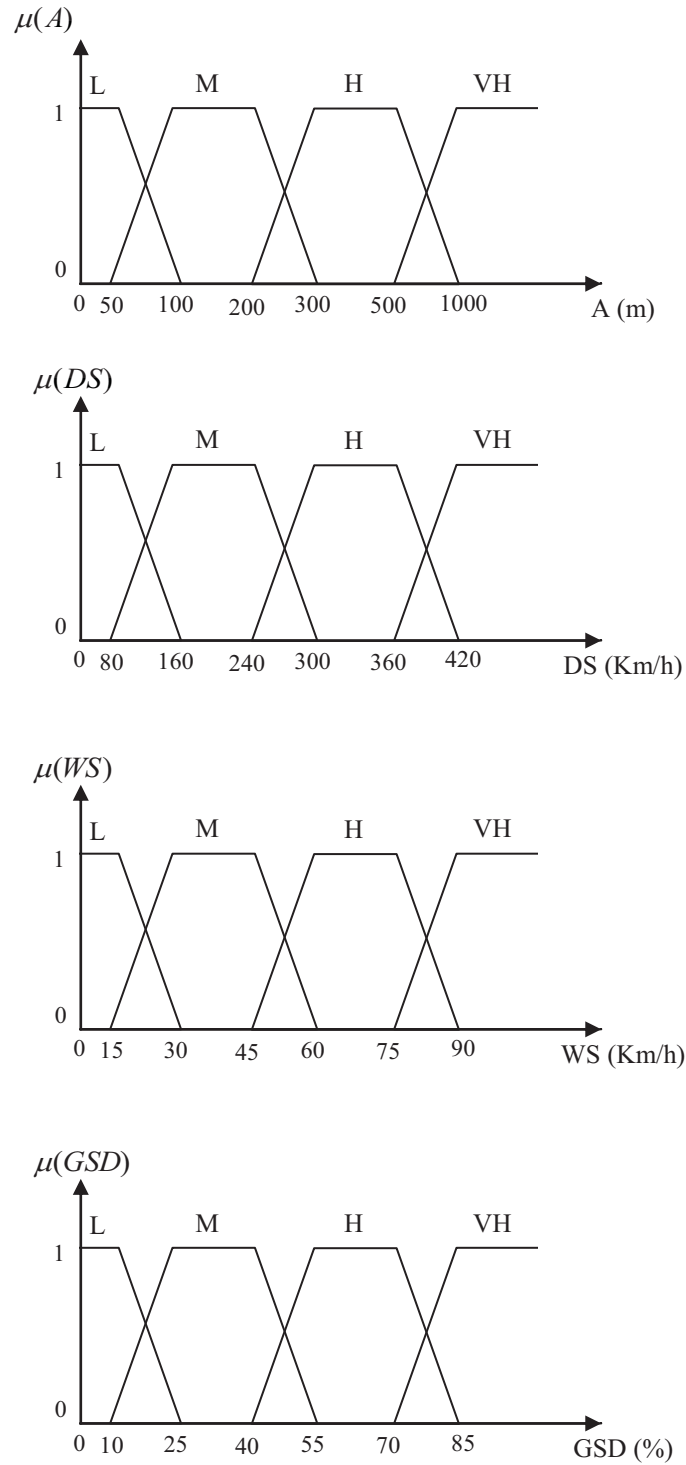


Figure 2. The suggested membership functions for the Fuzzification step. In each function four main fuzzy sets are considered (Low, Medium, High and Very high). The functions represent respectively: (the drone Altitude, the drone Speed, the Wind Speed and finally the Green Surface Density).

In our study, four different inputs are considered. Therefore, different partial tables have to be discussed in order to design the behavior of the system. For each partial table two inputs have to be pre-assigned to some pre-selected fuzzy sets. Thus, each cell in the table represents a suitable output for their assigned states and for a given states of the two other inputs. As an example, before filling out the Table 1 the following supposition will be formulated:

If (Drone Altitude = L) *and* (Drone Speed = L)

For the same state which has already assigned to the Drone Altitude, two new fuzzy states will be assigned to the Drone Speed. Thus, two new expressions will be considered:

If (Drone Altitude = L) *and* (Drone Speed = M)

If (Drone Altitude = L) *and* (Drone Speed = H or VH)

For each one of the last formulated expressions new table, which is describing the desired output for the given states of the inputs (Wind Speed and Green Surface Density), will be built. Six other expressions have also to be formulated to guarantee that all the possible cases of the spraying operation are taken in account, and for each expression one associated table will be also made.

If (Drone Altitude = M) *and* (Drone Speed = L)

If (Drone Altitude = M) *and* (Drone Speed = M)

If (Drone Altitude = M) *and* (Drone Speed = H or VH)

If (Drone Altitude = H or VH) *and* (Drone Speed = L)

If (Drone Altitude = H or VH) *and* (Drone Speed = M)

If (Drone Altitude = H or VH) *and* (Drone Speed = H or VH)

Thus, the entire behavior of the system becomes concluded in nine tables. In each table 16 output states are defined and capable of being modified at any time by the on-site staff to realize the desired performance of the system. From where we can say that Fuzzy Logic Controllers are easy to be designed and sufficiently flexible enough to be regulated depending on the changes that may arrive at any time to the environment of the system.

Table 1. *In this table, each cell describes a special suitable output for a given inputs (Wind Speed and Green Surface Density) and the pre-assigned states of (Drone Altitude and Drone Speed).*

| WS GSD | L | M | H | VH |
|-----------|---|---|----|----|
| L | L | M | H | VH |
| M | M | H | H | VH |
| H | H | H | VH | VH |
| VH | H | H | VH | VH |

2.3 Defuzzification step

In this final step, the decided quantity of the pesticides which has to be applied over the crop, will be converted from its fuzzy value (linguistic term) to its equivalent in crisp logic (digital value). To do that, the triangular membership function illustrated in Fig. 3 will be used. For some given inputs, if only one fuzzy rule in the inference engine was activated, only one fuzzy state with a special value between zero and one will be assigned to the output. Applying this last fuzzy value on its appropriate fuzzy set in the triangular membership function will be sufficient to determine the digital output value.

In other cases, if some inputs are considered like (Low Medium, Medium High or More High) two conditional rules in the inference engine may be activated. In this case, the theory of (center of gravity) will be employed to determine the output in crisp logic. The controlled output in the actual study is considered to be the radius of the spraying valve opening which is measured in millimeters as it is shown in Fig. 5. In our study, the head of the spraying machine is composed of one hundred valves.

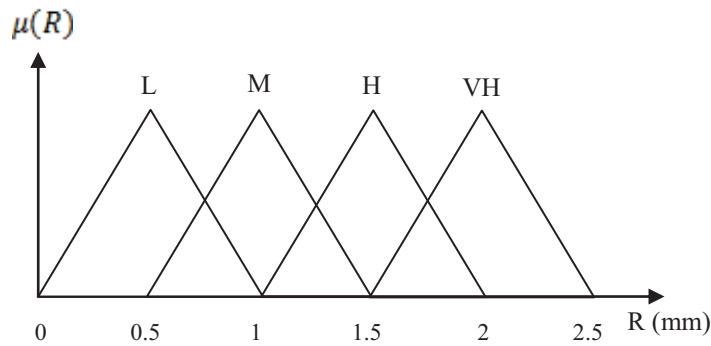
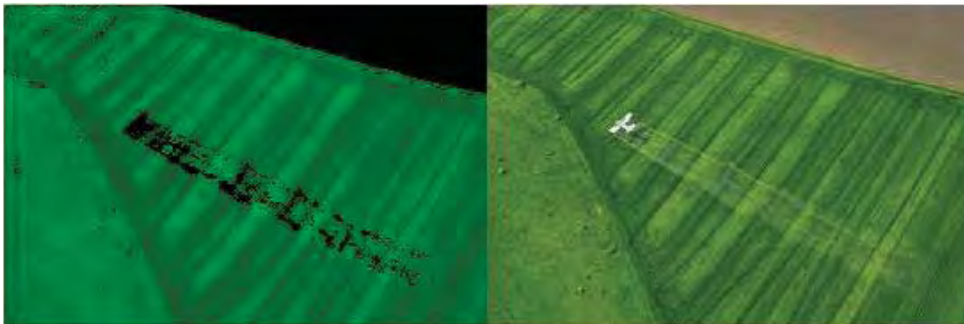


Figure 3. Triangular membership function is selected for the Defuzzification step. In this function four main fuzzy sets are considered (Low, Medium, High and Very high). The digital output is considered as a radius of spraying valve opening.

3. GREEN SURFACE DENSITY

A.



87.6891 %

B.



49.6843 %

Figure 4. Two captured images using Unmanned Aerial Vehicle. In the right side the original image, in the left side the extracted green component. In each case the percentage of Green Surface Density is illustrated.

The colored images which are taken by a mounted camera on the front of the drone represent in this study an important sensorial source about the crop over which the pesticides have to be applied. The percentage of the extracted green component from any given image will be called Green Surface Density which is one of our system inputs. To calculate this value, the data class of the RGB image has to be converted into scaled double as a first operation. Then for each pixel, the following condition has to be satisfied to be considered as a green pixel:

If (the green component of a pixel is $>$ Gr. Thresh.) *and*
 ((its red component is $<$ Min. Thresh. *and* its blue component is $<$ Max. Thresh.) *or*
 (its blue component is $<$ Min. Thresh. *and* its red component is $<$ Max. Thresh.))

Otherwise, it will be considered one black pixel as it is illustrated in Fig. 4.

In the last conditional expression three thresholds are defined: (Gr. Thresh) is a threshold of green component of any pixel. (Min. Thresh.) and (Max. Thresh.) are respectively the minimal and the maximal thresholds of red and blue components for any pixel. These last thresholds are experimental values, in our study they are selected as follows: (Gr. Thresh. = 0.3), (Min. Thresh. = 0.4) and (Max. Thresh. = 0.7). For some other applications we may try to evaluate Yellow Surface Density (YSD). In such a case, the last defined condition with its three associated thresholds has to be changed.

Then, Green Surface Density as a percentage will be calculated as follows:

$$GSD = \frac{n}{MN} * 100 \quad (1)$$

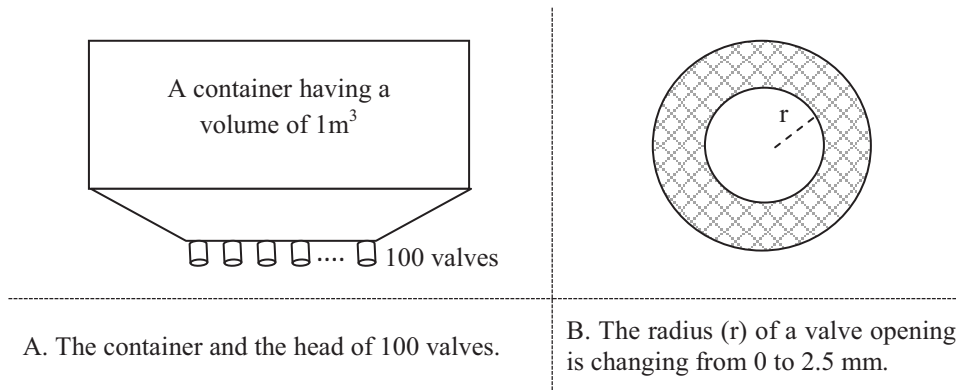
n: is a total number of green pixels

[M N]: are the rows and columns number of the colored treated image.

4. THE SPRAYING OPERATION

4.1 The spraying machine

The spraying machine represents the final actuator in this application, it is composed of a container having a volume of 1m^3 , and a head of 100 valves as it is shown in Fig. 5. The radius of a valve opening is supposed to be changeable between 0 and 2.5mm.



A. The container and the head of 100 valves.

B. The radius (r) of a valve opening is changing from 0 to 2.5 mm.

Figure 5. Simplified drawing for the spraying machine and the changeable radius of a valve opening.

The volume rate of liquid flow from a valve is given in the following equation:

$$\frac{V}{t} = Av \quad (2)$$

V: the volume of the container

A: the area of a valve opening ($\pi.r^2$)

v : the speed of the flowing liquid from a valve
 t : the time during which the total quantity of liquid was spread out the container

4.2 Experimental considerations

The speed of the flowing liquid from a valve is defined as: $v = 0.25$ m/s. The total area of one hundred opened valves with a constant radius of (2.5 mm) is: $A = 0.0020$ m².

Applying 1 m³ of pesticide, with a constant valves opening (without using the proposed approach in this paper) will be achieved during:

$$t = \frac{V}{Av} = 2000 \text{ s} \approx 33 \text{ minutes.}$$

Now, to compare the performance of the smart spraying operation which is presented in this paper with the classical method, the volume of pesticide which will be applied over the crop during the same last calculated time has to be determined. As an example: if the operation of pesticide spraying was divided for five supposed successive phases during which the valves opening radius was changing to take respectively the following values: (0.5mm, 1mm, 1.5mm, 2mm and 2.5mm) depending on some supposed inputs for each different phase. The main mission of covering each different region by the appropriate quantity of chemical materials depending on the density of its green surface and the other parameters will be achieved as it is illustrated in Fig. 6. And the total volume of the sprayed pesticide during the same 33 minutes will be only 0.44 m³ (440 lit). Thus, we are reducing the consumption of pesticide by 56 % comparing with the classical method. In addition, thanks to its capability of calculating the GSD factor depending on some techniques of digital image processing, the proposed fuzzy logic controller is capable of solving the fundamental problem of applying the chemical materials outside the crop limits.

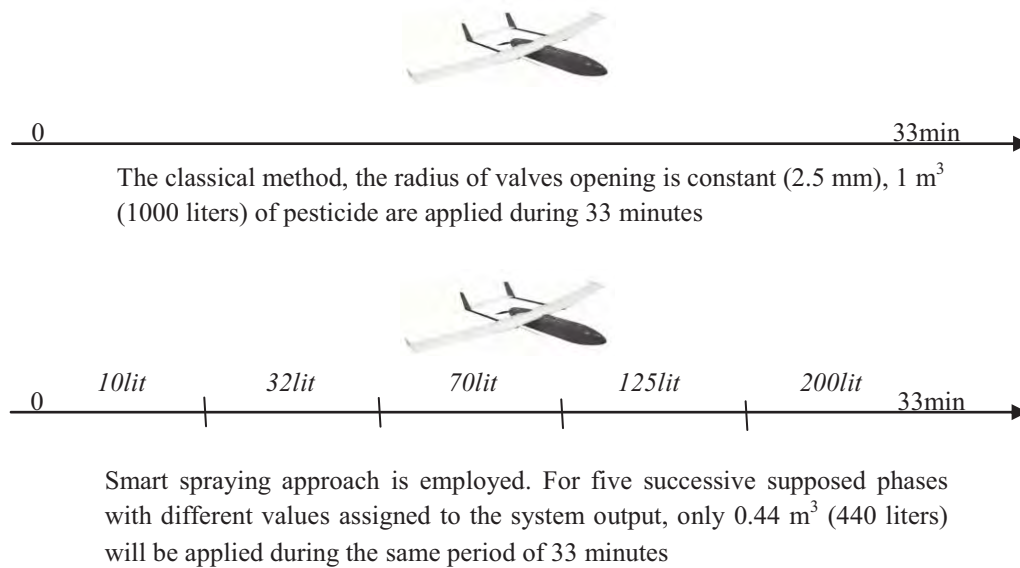


Figure 6. Comparing the performance between the classical spraying method and the smart approach proposed in this paper.

5. CONCLUSION AND PERSPECTIVES

In this paper, we present our study to build a pesticide spraying system that depends on some advanced techniques of fuzzy logic and digital image processing in order to reduce the quantity of applied chemical materials over the green crops. This smart system may represent an economic and ecologic choice for the farmers. As it may solve some of their practical problems like spraying outside the crop limits which may increase the risk of having some negatively affected or polluted areas. The proposed approach is capable of achieving the desired performance of covering

precisely the needed areas by the appropriate quantities of pesticide with reducing the total volume which was needed to cover the same surface using the classical spraying method. All the discussed and illustrated functions in this paper are already built and tested in MATLAB. The practical and experimental phases of this project will be started nearly, as soon as the drone, the sprayer and the sensors are purchased and assembled together as a final product in order to realize our objective of having an autonomous smart agricultural drone.

6. REFERENCES

- [1] G. You, X. Wang, S. Yang, D. Wang, "Precision Spraying System of Crops Disease Stress Based on Acoustic Emission", JOURNAL OF COMPUTERS, VOL. 6, NO. 4, APRIL 2011
- [2] J. W. Miller, "Report on the development and operation of an UAV for an experiment on unmanned application of pesticides" Youngstown, Ohio: AFRL, USAF, 2005.
- [3] R. Sugiura, K. Ishii, N. Noguchi, "Development of Monitoring System to Support Operations of an Unmanned Helicopter". Paper No. 05-1019, American Society of Agricultural and Biological Engineers. Presented at Tampa, FL USA. 2005.
- [4] S. Yan, "Study on the System of Spraying Rate Varied by Pressure of Liquid Chemical Application", Chinese Agriculture University, 2004.
- [5] Y. Huang, W.C. Hoffmann, Y. Lan, W. Wu, B.K. Fritz, "Development of a Spray System for an Unmanned Aerial Vehicle Platform". Applied Engineering in Agriculture, 25:803-809. 2009.
- [6] Y. Ru, H. Zhou, Q. Fan, X. Wu, "Design and Investigation of Ultra-low Volume Centrifugal Spraying System on Aerial Plant Protection". Paper No. 11-10663, American Society of Agricultural and Biological Engineers. Presented at Louisville, KY USA. 2011
- [7] Z. Wang, Y. Lan, W. C. Hoffmann, Y. Wang, Y. Zheng, "Low Altitude and Multiple Helicopter Formation in Precision Agriculture". Paper No. 13-1618681, American Society of Agricultural and Biological Engineers. Presented at Kansas City, MO, USA. 2013.

Smart template matching algorithm using hill-climbing search strategy

Abdellatif BABA

Mechatronics department, University of Turkish Aeronautical Association,
Bahçekapı Quarter Okul Street, Etimesgut, ANKARA, Turkey

ababa@thk.edu.tr

ABSTRACT

More intelligent and faster version of template matching algorithm is presented in this paper. This advanced method employs the Hill-Climbing strategy to achieve local search around limited number of randomly sampled pixels, in order to determine the best correlation between two digital models. Theoretical explanations and practical implementations will be successively illustrated.

Keywords – Template matching, hill climbing, random sampling.

1. INTRODUCTION

Template matching is one of the most widely known algorithms, which is normally used to verify the correlation state between an indexed digital model (template) and a part of a given image. In spite of its good competence, its slow performance represents its main drawback. Our aim is to produce a new and faster version of that algorithm in order to make it more suitable for autonomous and embedded systems, like in robotics applications. In the next paragraph we discuss some related works. In the 3rd paragraph, the classical version of Template Matching Algorithm is illustrated. Then, we explain quickly in the 4th paragraph the initial mechanism of “Hill-Climbing” search strategy. Then, all the details of our new developed algorithm are explained in the 5th paragraph. To prove the efficiency of our algorithm, experimental results are shown in the last section. Then, we conclude with perspectives.

2. RELATED WORKS

Nixon et al. [11] present the classical and famous version of template matching algorithm, which is a time consuming method, as it becomes useless in the case of having any rotational angle or scale difference between the two compared models. Other approaches [7, 14] propose working in frequency domain, but they have some disadvantages in digital implementation, and lack of reliability to find the location of the best matching. Several approaches [10, 18] propose configuration similarity retrieving in query processing. Papadias [5] employ hill-climbing algorithm as a search technique to satisfy each proposed query. The three last mentioned approaches may be classified according to the size of database images and the type of query variables (static or dynamic) which may not be simple to be defined, as they need a lot of constraints. Our approach proposed in this paper doesn't call the problem of similarity retrieving in query processing, but it employs random sampling and hill-climbing local strategy as a smart technique to verify the best correlation between an indexed model and a current image. Our approach is applied in spatial domain; it aims to accelerate the execution time of the classical algorithm. In addition, we employ some techniques to overcome problems like angular rotation or scale variance as it will be seen in the 5th paragraph.

3. TEMPLATE MATCHING (The CLASSICAL ALGORITHM)

The mechanism of this algorithm is shown in Fig. (1), where the template of (n) pixels has to scan the entire image. In each new location, the error has to be calculated as illustrated in Eq. (1), this determined error has to equal zero if the best correlation is detected.

$$Error_{y,x} = \sum_1^n (img_{y\pm i, x\pm j} - Temp_{y\pm i, x\pm j})^2 \quad (1)$$

(i, j) : two integers represent the displacement form the position (x, y) where the template is actually located.

For an image of [M N] dimensions and a template of (n) pixels, one matching case may be discovered using this classical algorithm by applying (M*N*n) subtraction operations, which means a heavy calculation load. On the other hand, this algorithm becomes incapable of verifying any correlation state if there is any small rotational angle or any scale difference between the template and the image.

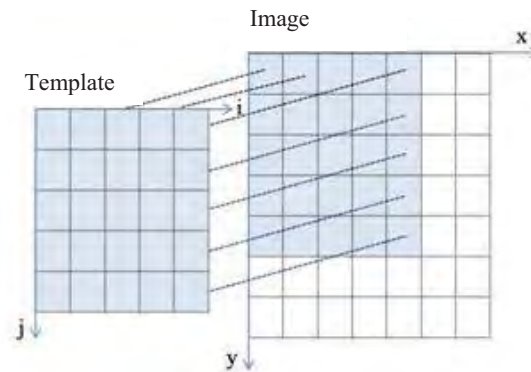


Figure 1. *The mechanism of classical algorithm*

4. HILL-CLIMBING ALGORITHM

This algorithm is a loop that starts from initial state (node) which is selected randomly in the state space, and continuously moves in the direction of increasing or decreasing value of the objective function of the neighbored states. The aim is to reach the state for which the value of the objective function represents global maximum or global minimum [1, 9, and 16]. For some cases when the objective function reaches local maximum or local minimum, the algorithm stops its searching activity without getting the desired solution, such case has to be distinguished and new random state has to be selected from the state space to start a new iteration.

5. THE IMPROVED ALGORITHM

The principle of our improved algorithm which is presented in this paragraph is faster, more intelligent and capable of solving problems like scale and angular variance. The two successive tables (I) and (II) depict clearly its steps. At first, limited number of pixels in the original image has to be randomly sampled. To affect all the image regions by the sampling process, we propose to divide the image into limited number of virtual regions, from each region the same number of samples has to be randomly taken as it is shown in Fig. (2).

In the second step, template matching has to be verified in the last selected locations. But, instead of using the classical Eq. (1), the average of template pixels intensities has to be calculated as in Eq. (2). Then, in each location this last average has to be subtracted, Eq. (5), from the average of image pixels intensities, which is already determined in Eq. (3). This last calculated error represents the objective function which has to be minimal in the position of the best correlation. Therefore, this algorithm doesn't require the similar pixels in the two arrays (the template and the sub-image) to be faced in the same order to prove one correlation case. But, it subtracts simply two averages of two faced areas, as illustrated in Fig. (3). Therefore, verifying the best correlation between the two digital models, even with the existence of some rotational variance, needs only to detect the location of the minimal error, because it is impossible to have zero error in such cases.

Table 1. *The developed algorithm.*

$TempAvrg = \sum_{i=1}^{T_{Col}} \sum_{j=1}^{T_{Row}} Temp_{j,i} / n$
 Initial state has to be randomly drawn

 For the initial state : $i = 0$
 Calculate $ImgAvrg(y, x)$
 (Call Correlation Check Procedure)

for ($j = 1:N$)

 for ($i = 1:n$) neighbored suggested states
 Calculate $ImgAvrg(y, x)$
 (Call Correlation Check Procedure)
 end for

 Determine the Minimal Error

 if $Minimal\ Error > Error_{i=0}(y, x)$
 Local Minimum case is distinguished, new initial
 random sample has to be drawn
 else
 The state that has the Minimal Error is selected. New
 (n) locations will be randomly drawn for the next
 iteration.
 end if
end for

Table 2. *The correlation check procedure*

$Error_i(y, x) = abs(ImgAvrg(y, x) - TempAvrg)$
if $Error_i(y, x) == 0$
 The Solution (Global Minimum) is found
 break
end if

$$TempAvrg = \sum_{i=1}^{T_{col}} \sum_{j=1}^{T_{row}} Temp_{j,i} / n \quad (2)$$

(n) : the total number of template pixels

(T_{col}, T_{row}): represent respectively the number of template columns and rows.

$$imgAvrg = \sum_{i=1-cx}^{cx-1} \sum_{j=1-cy}^{cy-1} img(y+j, x+i) / n \quad (3)$$

(n) : the total number of the sub-image pixels

(cx, cy) : are respectively the half number of template columns and rows, they are calculated in the following equations Eqs. (4).

$$cx = \text{floor}\left(\frac{Tcol}{2}\right) + 1, \quad cy = \text{floor}\left(\frac{Tcol}{2}\right) + 1 \quad (4)$$

(floor) : is a MATLAB function that rounds the real number to its nearest integer towards minus infinity.

$$\text{Error}(y, x) = \text{abs}(\text{imgAvg}(y, x) - \text{TempAvg}) \quad (5)$$

(abs) : is a MATLAB function that determines the absolute value of an element.

Then, the calculated error in each randomly selected location has to be evaluated. If it is bigger than one pre-selected threshold, new random sample has to be taken from its neighborhood. Else, the template has to achieve its local search using Hill-Climbing search strategy, which is shown in the Fig. (4). Where, the template has to select one of eight possible locations around it. To decide for which one it has to move the error has to be calculated in each location and the minimal one will be the best choice. From this last selected location, the template has to go farther in order to check the best correlation state which may exist in any different location in the image. To reduce the execution time as much as possible, the direction of each current step (during the template movement in its trajectory) will influence the direction of the following step as it is illustrated using the blue arrows in the Fig. (4). As an example, if the algorithm has already selected the north east as a movement direction to arrive its best location, the most favorite directions for the following step will be only one of three possibilities (north, east and north east), and so on.

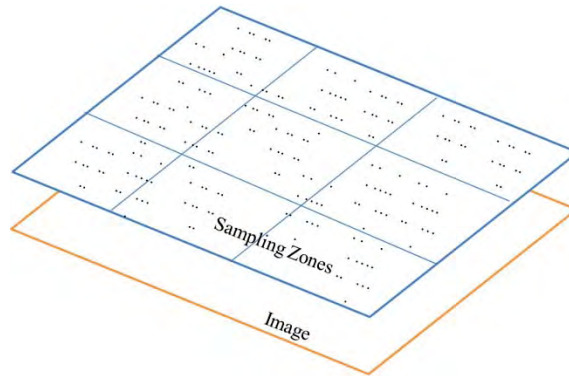


Figure 2. Up to the image, virtual mask is used to guarantee the best distribution of the randomly selected samples.

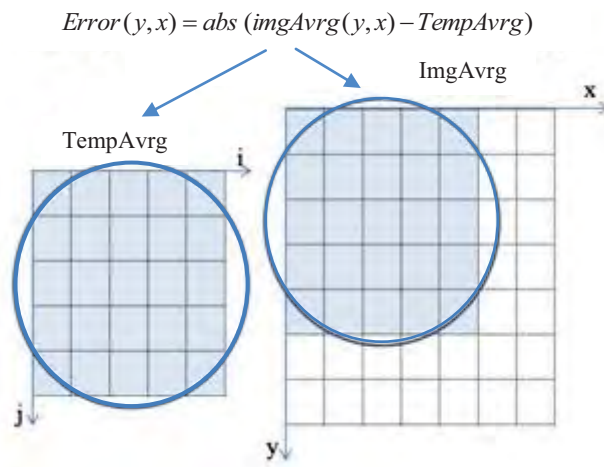


Figure 3. In one randomly selected location, the two averages are subtracted.

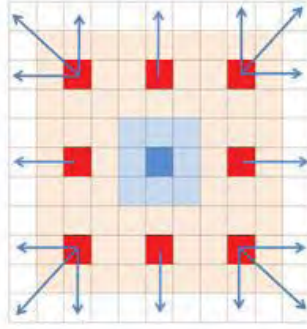


Figure 4. Light blue cells represent the template pixels, and the template center is in dark blue. Red cells represent the eight possible neighbored locations. Blue arrows, points to new locations will be selected randomly with keeping the influence of the last step direction.

5. IMPLEMENTATION AND RESULTS

5.1 Efficiency verification of the developed algorithm

We apply using MATLAB this developed algorithm on a gray scale image illustrated in the Fig. (5), which has the size of [500 700]. The template is illustrated in the same figure as a part of this image, its size is [51 51]. The used microprocessor in this implementation is (Intel Core i5, 2.30 GHz). In this experiment we take 1500 samples from 9 different virtual zones of the image. Thus, we are speaking about 13500 samples. This last operation takes place, at most, in 3 seconds. This last time will be in the range of milliseconds if the algorithm was applied using C or C++. Applying the same experiment using MATLAB and the same microprocessor, but with the classical algorithm consumes nearly 10 seconds. So, we can conclude that the new algorithm is three times faster than the classical one.

5.2 Places Recognition

In this experiment, a mobile robot localized in (x_r, y_r, θ_r) takes an image in a moment (t) as it is shown in the Fig. (6). The most important part of the image that contains the distinctive features was selected and indexed in the robot data base, this selected part will be considered later as a template.

If in another moment $(t+\delta t)$, the same robot takes another image for the same scene, as in the Fig. (7), but from a new position $(x_r + \delta_x, y_r + \delta_y, \theta_r + \delta_\theta)$ and a new point of view. In this case, scale and angular variance between the indexed template and the current image exist. Therefore, the desired correlation hasn't any chance to be detected using the classical algorithm. But our improved version determines the position where the least local minimum was detected in the image, as it is marked by the small red cross in the Fig. (7).

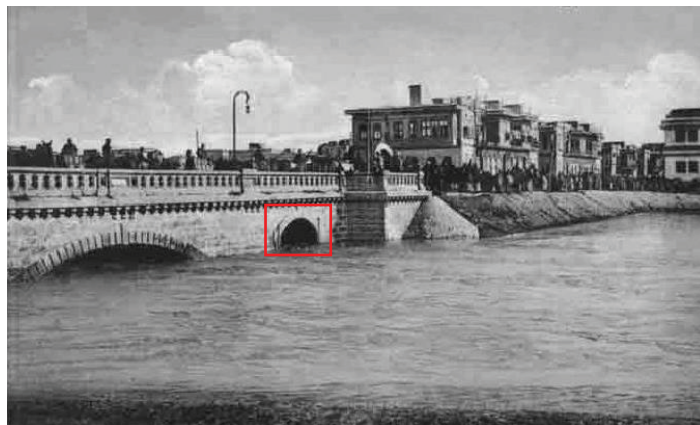


Figure 5. Gray scale image of size 500x700. The selected template of size 51x51 is surrounded by a red square



Figure 6. *The first image is taken by mobile robot in the position (x_r, y_r, θ_r) in a moment (t) , the template which is surrounded by red square, is selected as a part of the image that contains some important features*



Figure 7. *Another image is taken by the same mobile robot in another new position $(x_r + \delta_x, y_r + \delta_y, \theta_r + \delta_\theta)$ in a moment $(t + \delta t)$; the template matching is realized by searching the least local minimum which was detected in one point which is marked by a small red cross.*

6. CONCLUSION

In this paper, we propose the employment of Hill-Climbing search strategy to improve the functionality of Template Matching algorithm in its classical version. The main development, on which we concentrate here, is to reduce the execution time of the algorithm with keeping its efficiency. Some technical ideas are suggested and discussed to solve some practical problems like scale and angular variance between the image and the template. The new algorithm is explained in all its steps, and some experimental results are illustrated whether to prove its better performance in comparison to the classical version or to exploit it in some applications like places recognition which is useful for mobile robot localization.

To realize the best version of this algorithm, we intend to use parallel programming strategies to achieve simultaneous random sampling and the associated local search in several locations in the image. This idea may be helpful to reduce more and more its execution time to arrive an optimal performance which is required for some applications like landmarks recognition depending on successive indexed images to localize an autonomous unmanned aerial vehicle.

7. REFERENCES

- [1] A. Cawsey. The Essence of Artificial Intelligence. Prentice-Hall, 1998.
- [2] A. Leon-Garcia, "Probability and Random Processes for Electrical Engineering, 2nd ed. Addison-Wesley, Reading, MA. 1994
- [3] D. Hanselman, and, B. R. Littlefield, "Mastering MATLAB 6", Prentice Hall, Upper Saddle River, NJ. 2001.
- [4] D. J. Lowe "Distinctive Image Features from Scale-Invariant Keypoints", Computer Science Department University of British Columbia, 2004.
- [5] D. Papadias, "Hill Climbing Algorithms for Content-Based Retrieval of Similar Configurations", Proceedings of the ACM Conference on Information Retrieval (SIGIR), Athens, July 24-28, 2000.
- [6] G. Taylor, and, L. Kleeman "Visual Perception and Robotic Manipulation", ISBN 10 3-540-33454-8, 2006.
- [7] J. Altman, and, H. J. P. Reitbock "Fast correlation method for scale and translation invariant pattern recognition, IEEE Trans. On PAMI, 6(1), pp. 46-57, 1984.
- [8] J. C. Russ, "The Image Processing Handbook", 3rd ed. CRC Press, Boca Raton, FL, 1999.
- [9] K. Sullivan and S. Jacobson. A Convergence Analysis of Generalized Hill-climbing Algorithms. IEEE Transactions on Automatic Control, 46(8):1288–1293, 2001.
- [10] M. Nabil, A. Ngu, J. Shepherd "Picture Similarity Retrieval using 2d Projection Interval Representation". IEEE TKDE, 8(4), 1996.
- [11] M. Nixon, and, A. Aguada "Feature Extraction & Image Processing", ISBN 0750650788, 2002
- [12] P. Jensfelt "Approaches to Mobile Robot Localization in Indoor Environment", ISBN 91-7283-135-9, 2001.
- [13] R. C. Gonzalez, R. Woods, and S. Eddins "Digital Image Processing using Matlab", ISBN 0-13-008519-7, 2004.
- [14] R. N. Bracewell "The fourier transform and its application", 2nd edition, McGraw-Hill Book Co, Singapore, 1986.
- [15] R. M. Haralick, and L. G. Shapiro, "Computer And Robot Vision", vols. 1 & 2, Addison-Wesley, Reading, MA. 1992.
- [16] S. Russell, and ,P. Norvig "Artificial Intelligence, A Modern Approach" , ISBN 0-13-080302-2, 2003.
- [17] S. Thrun, W. Burgard, D. Fox "Probabilistic Robotics", ISBN 0-262-20162-3, 2005
- [18] V. Gudivada, V. Raghavan "Design and evaluation of algorithms for image retrieval by spatial similarity". ACM Transactions on Information Systems, 13(1), 115-144, 1995.

Analyzing the effect of the slit depth and width on the electrical performance of a squirrel cage induction motor

A G Yetgin¹, M Turan², E Unlukaya³

^{1,3}Electrical and Electronics, Dumlupınar University,
Tavsanlı Street, Central, Kutahya, TURKEY

²Electrical and Electronics, Sakarya University,
Esentepe Street, Serdivan, Sakarya, TURKEY

¹*gokhan.yetgin@dpu.edu.tr*, ²*turan@sakarya.edu.tr*, ³*efe.unlukaya4@ogr.dpu.edu.tr*

ABSTRACT

In this study, a new design was suggested in order to improve the performance of induction motors. In these new designs, slits were applied in the middle of stator and rotor tooth. In these slitted models, the depth and width in the 56 different slitted motor models were optimized by Finite Element Method Magnetics (FEMM) software by using Finite Elements Method (FEM). What value the depth and width of optimum slit should be was determined in order to obtain nominal torque in the new motor models created with the proposed slitted structure, and how the depth and width of slit could affect the torques of motor was demonstrated. In modeling, polyphase, 3 kW, squirrel cage induction motor was used. As a result of modeling, increase in the coupling flux provided due to slits, 0.338%, 2.085% and 4% increase in starting, pull - out and nominal torque values respectively for 15 mm slit depth and 0.1 mm slit width.

Keywords – Induction motor, slit optimization, speed - torque characteristic, finite element method

1. INTRODUCTION

Induction motors which have high efficiency, easy control with the development of power electronics, rugged, cheaper due to mass production etc. are the most commonly used motor in industrial application. Especially torque - speed characteristic of induction motor is very important role for a lot of applications [1]. Many different approaches have been applied in order to improve the performance of induction motor especially in the researches focused on the energy saving issues. It is possible to collect these approaches under 4 titles [2]. One of them is improve the performance with the works on slot shapes by using design software [3]. Second is improve the performance with optimization methods such as Finite Elements Method, Artificial Neural Networks (ANN), Fuzzy Logic (FL) and Genetic Algorithms (GA) [4], [5]. The other is improve the performance by developing the materials used [6]. Last is improve the performance by using magnetic barriers, cut - outs and slits [7].

In this study, slits were applied between the stator and rotor teeth of induction motor in the proposed design. How the torques of motor was changed in different slit depths and widths and optimum slit depth and width were determined for maximum motor torques.

2. MAGNETIC BARRIERS, CUT - OUTS and SLITS REVIEW

Recently, to improve induction motor performance through alternative core design some other approaches have been used such as magnetic barriers, cut - outs and slits. Those designs have been applied especially on solid rotor induction motor [8], hybrid excitation synchronous motors [9], switched reluctance motor [10] and interior permanent magnet motor [11]. In the literature, until, there is no usage of slitted design for laminated core squirrel cage induction motors except for

author papers [12], [13]. A few studies using slitted design are examined below for different motors.

Aho, Nerg and Pyrhonen, in their studies, examined the changes of magnetic rotor by adding slits to the rotor structure of solid rotor induction motor. They showed that when the number of slit was increased from 28 to 36, the electromagnetic torque could be improved by 6% and power factor could a little be increased. Moreover, they stated that when the number of rotor slit numbers were increased saturation occurs between slits and so it worsened the electromagnetic performance. They showed that as the number of slit was increased, iron losses were increased [7]. Pyrhonen et al., investigated the torque values changes by slit depth for solid rotor induction motor. They showed that the best torque is reached when the slit depth is 60% of the rotor radius [8]. Li et al., suggested slitted structure model in order to prevent the armature reaction in hybrid excitation synchronous machine. They stated that the effect of armature reaction could be reduced by making flux line longer with slitted structure [9]. Hybrid excitation synchronous machine with slitted construction is given Figure 1.

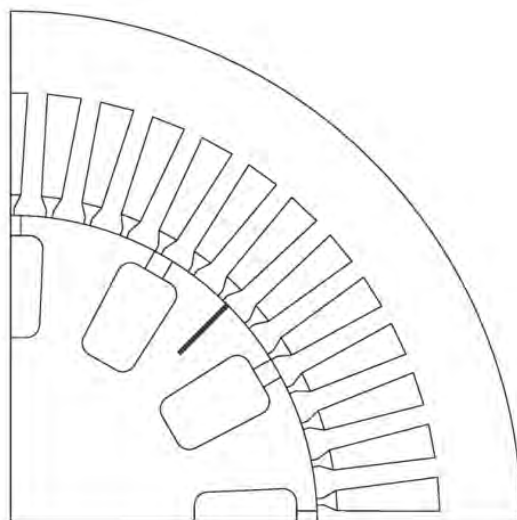


Figure 1. *Hybrid excitation synchronous machine with slitted construction.*

In their study, Chan and Hamid examined the current, torque and magnetic flux density changes by slitting in various numbers in the rotor structure of switched reluctance motor by using FEM. In the study, they stated that a flat topped wave form was obtained in both flux and torque graphics for five slitted models and the output power increased by 16% (without increase in top value of current). Moreover, they stated that the saturation in rotor could be controlled by changing the slit number [10]. Switched reluctance motor with slitted construction is given in Figure 2.

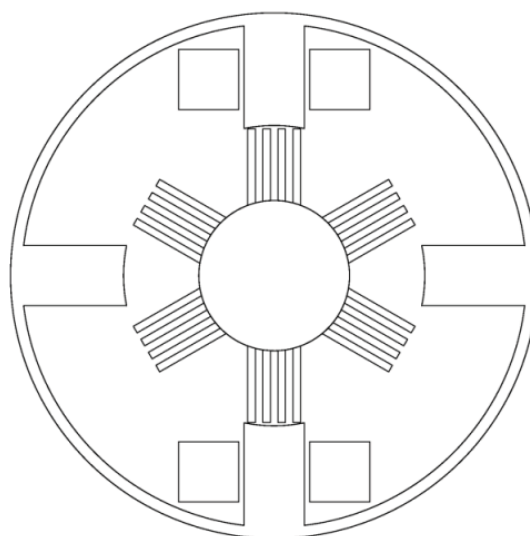


Figure 2. *Switched reluctance motor with slitted construction.*

Soong and Ertugrul compared synchronous reluctance and Interior Permanent Magnet (IPM) motors versus 2.2 kW induction machine. 5 prototype rotors, one of them is standard squirrel cage rotor, the other two are axially laminated synchronous and an IPM rotor which is given in Figure 3, the last are four barriers and five barriers designs were construction and tested [11].

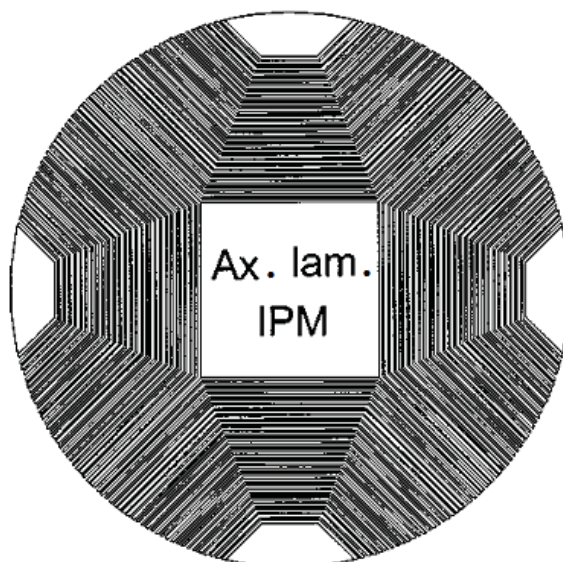


Figure 3. Axially laminated interior permanent magnet rotor with magnetic barriers.

Yetgin and Turan, in their study, the motor models reformed with the proposed shape design were analyzed with FEMM software program that uses finite elements method. It was found at the end of the analyses that when the optimum slit width value was 0.10 mm, motor efficiency gave better results compared with other slit width values. The increase in slit width caused saturations in motor teeth and thus caused to worsen the motor efficiency. It was decided at the end of different modeling that the depth of the slits used in the proposed slitted motor models should be almost the same height as stator and rotor height. Optimum slit value was determined as 15.00 mm. At the end of the modeling, 1.869% improvement was obtained in the motor models that had 15.00 mm slit depth and 0.10 mm slit width compared with the efficiency values of reference model at nominal operation point [12]. Yetgin et al., by opening slit in the stator and rotor tooth of the induction motor and reduced the zigzag fluxes, which arise in the air gap of the motor, by 6.123%, ensuring a 2.041% improvement in the coupling flux. With the point measurements in terms of their flux densities, it was illustrated that less strain is present compared to the reference motor [13]. Aho, Nerg and Pyrhonen investigated the effects of the slit depth of solid rotor induction motor which had a slitted structure on the performance in Figure 4. They indicated that the slitted rotor created a better field distribution, but reduced the mechanical strength of rotor and it was difficult to analyze the slitted motor structure analytically. They suggested that the deeper the slit, the more torque was obtained, and a huge increase was obtained in the power factor value at the end of their study. They stated that the mechanical brittleness of rotor increased in the slitted forms which had bigger depth, and in order to prevent it, 1 short and 1 long slitted form could be used. Moreover, they stated that the depth of slit should be the half of the radius of the rotor [14]. Solid rotor induction motor with slitted construction is given in Figure 4.

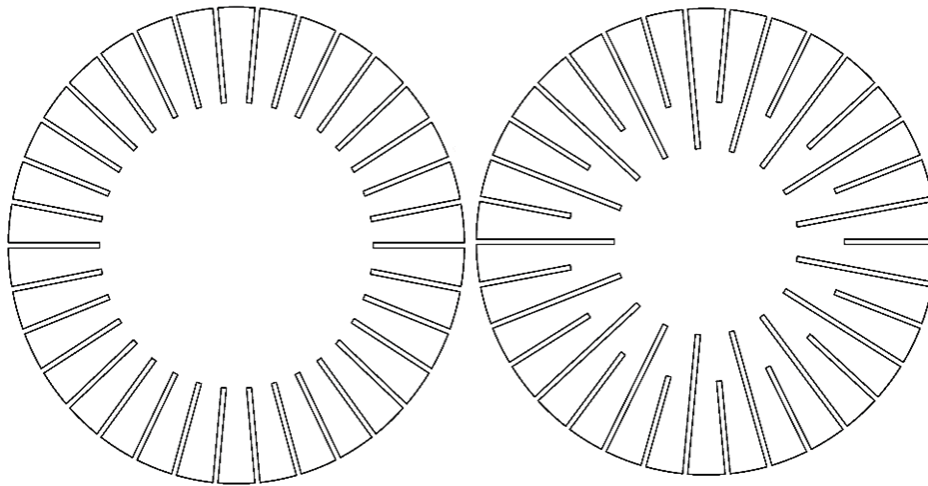


Figure 4. *Solid rotor induction motor with slitted construction.*

Matsuo and Lipo presented three generation type synchronous reluctance motor which has barriers, segmental rotor and axially laminated anisotropic rotor. Synchronous reluctance motor with magnetic barrier is shown in Figure 5 [15].

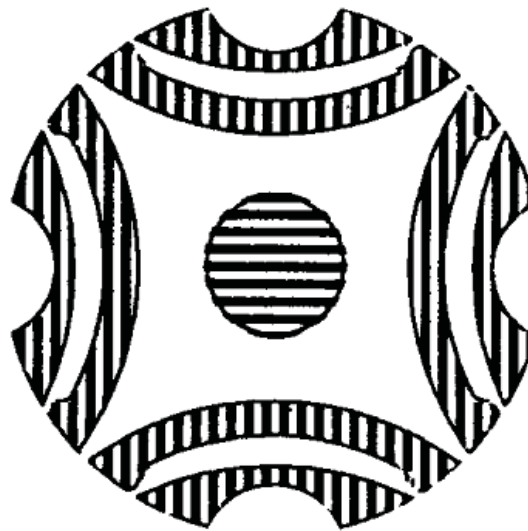


Figure 5. *Synchronous reluctance motor with magnetic barrier.*

Zaim, in his study, calculated with FEM how motor performance changes according to the slit width, depth and number of the solid rotor induction motor. He showed that the more the slit depth increased, the more the torque increased at certain level and then it decreased. Moreover, he determined that the torque decreased as the width of slit increased. He stated that while torque first increased rapidly and then decreased as the number of slit increased in narrower slitted forms, in wider slits, the torque first increased and rapidly decreased. Furthermore, in that study, it was stated that rotor flux lines proceeded towards the inner sides of rotor by choosing the suitable slit depth and width [16]. In the study Nashiki et al., they placed slits on the rotor in order to decrease the torque oscillation and create magnetic flow in the direction of axis. They examined torque oscillations for the situations that had and did not have rotor skewing in slitted form. They stated that when the number of slit was low, torque oscillations could increase [17]. In his study, Zaim evaluated the performance of solid rotor induction motor that has various power, slit number, depth and width by using FEM. He showed that while the increase in slit width increased the torque of motor in the models that have low number of slit, it decreased the torque of motor in the models that have high number of slit. He also stated that when the slit depth was increased, the torque first increased and then it decreased [18]. Dorairaj and Krishnamurthy presented in their study the performance of the polyphase induction motor with slitted ferromagnetic rotor. Their experimental studies covered the influence of physical dimensions of slits and their number on the rotor performance, effects of inserting a conducting, nonmagnetic material like copper in the slits, and the performance of these rotors terminated with and without copper end rings [19]. Jagiela and

Garbiec evaluated the best length of the induction motor which has slitted solid rotor. They aimed to obtain maximum efficiency. If the motor's length of end region is increased from 0 to 12 mm, the motor efficiency is reducing from 0.58 to 0.53, also torque value is reduce from 0.34 to 0.21 [20].

In this study, a new core design is developed in order to improve the torque - speed characteristic of induction motor. In the core design, stator and rotor slot shapes, air gap length, stator outside diameter and shaft diameter etc. values are not changed. Starting, pull - out and nominal torque values gained from reference motor model and proposed slitted motor models and the change of torque - speed characteristics are given in comparison. Induction motor design is getting improved toward to excellence during the century. Proposed slitted tooth core design in the paper, not only improves the torque - speed characteristics of the induction motor, but also improves some other performance parameters which are discussed in the paper. Energy efficiency and quality matters are important issues in the millennium. The proposed design will be applicable since it offers a few points of improvements in the journey of excellency for the high end induction motors.

3. FINITE ELEMENT METHOD

The FEM is the most useful program for electrical engineering and motor designer. Its advantage is handle non - linear problems, time dependent and non-uniform geometry. The FEM process starts with the creation of a geometrical model and input the motor parameters which are problem type, frequency, depth, solver precision, boundary conditions, winding values etc. Later, mesh is created into small triangle. After mesh, the problem is solver and desired parameters are obtained [21].

There are many studies in the literature using FEMM program [22]-[26] and the results obtained from comparing the experimental and simulations are supported by each other. When FEMM simulation and experimental study were compared by [27], error rate was found as 0.065%. In this study, the formulas which used in FEMM models are taken from reference [28].

4. REFERENCE and PROPOSED SLITTED CONSTRUCTION MOTOR MODELS

New tooth geometry is proposed in order to improve torque - speed characteristic which is the most important performance criterion in induction motor without using higher grade lamination sheets. By proposed slitted construction, it is aimed to overcome rotor reaction effects by decreasing quadrature axis flux such as leakage and zigzag flux in the air gap and, to increase d axis coupling flux which is the major parameter of induced torque. Thus, induction motor performance will be improved by using slitted core design. Nameplate of reference motor is given in Table 1.

Table 1. Parameters of the induction motor.

| Motor Parameters | Symbol | Value |
|---------------------------|-----------|-------|
| Nominal Power [kW] | P_n | 3 |
| Nominal Voltage [V] [L-L] | U_l | 380 |
| Nominal Current [A] [L-L] | I_n | 6.63 |
| Frequency [Hz] | f_l | 50 |
| Power Factor | $cos\phi$ | 0.88 |
| Nominal Efficiency [%] | η | 78 |
| Rotor Speed [rpm] | n_r | 2844 |

Reference motor model whose analysis were carried out with FEMM software package is shown in Figure 6 and the motor model on which the proposed slitted structure was applied are shown in Figure 7a and 7b. The results and graphics are given for nominal operating point.

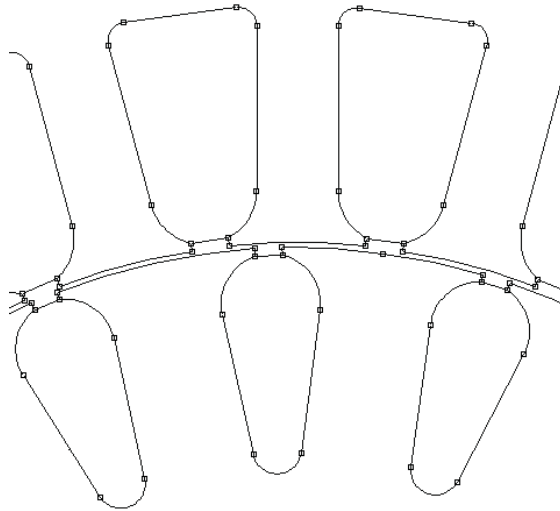


Figure 6. *The reference motor (R.M.) model.*

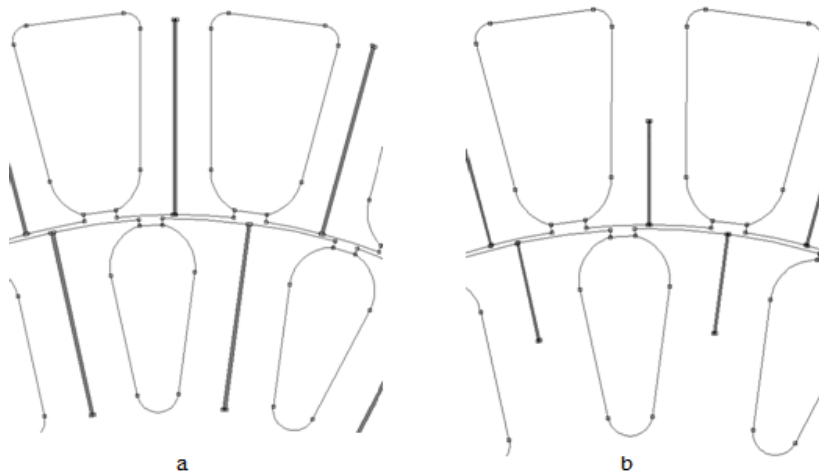


Figure 7. *a) Motor model which has 15 mm slit depth, 0.25 mm slit width b) Motor model which has 7.5 mm slit depth, 0.15 mm slit width.*

4.1 Advantages and Disadvantages of Proposed Slitted Construction Motor

The advantages and disadvantages of proposed slitted construction according to reference motor are [2]:

- i). The more effective using of the magnetic flux for all the working areas: reduction of saturation, properly distribution of flux at the teeth, reduction of rotor reaction
- ii). Decreasing additional losses: a properly flux distribution
- iii). The reduction of loss values: loss values improve due to decreasing saturation and iron losses
- iv). The reduction of slip: decreasing rotor copper loss
- v). The improvement of efficiency: the efficiency increases because of decreasing losses
- vi). The improvement of performance at the nominal operating points
- vii). The improvement of leakage reactance: the improvement of torque - speed characteristic
- viii). For small power motors, the initial construction cost is slightly high owing to cut off the core with laser.
- ix). For big power motors, there will be no extra cost due to the new mold to be created for mass production.

4.2 The Opening of the Slits

In proposed design, slits on the teeth of rotor and stator can be machined by laser cutting, wire cut EDM or water jet cutting machines. Sheets for small horsepower motors can be cut one by one or multiple at once up to 20 mm thicknesses. In application of proposed slitted design to high power motors, it is more convenient to use die cutting technics. Since stamping die mold is manufactured once, there will be no additional costs for high power motors in mass production.

4.3 Determine of Slit's Depth and Width

Slits were applied in the middle of both stator and rotor teeth of rotor squirrel-cage induction motor. In the models, slit depth were chosen from 6.25 mm to 23.00 mm. The slit widths were the models changing from 0.09 mm to 2.50 mm. For each slit depth models having slit width were created. An optimization work carried out on 56 different designs was presented in Table 2. Motor models having different slit depth and width were shown as x. Slit depth and width values are chosen to be made by laser cutting.

Table 2. Motor models having different slit depth and width.

| Width of Slit [mm] | Depth of Slit [mm] | | | | | | | | | |
|--------------------|--------------------|------|-------|-------|-------|--------|-------|-------|-------|-------|
| | 6.25 | 7.50 | 11.25 | 13.12 | 15.00 | 16.875 | 18.75 | 19.60 | 21.00 | 23.00 |
| 0.09 | | X | X | X | X | X | X | | | |
| 0.10 | X | X | X | X | X | X | X | X | X | X |
| 0.15 | X | X | X | X | X | X | X | X | X | X |
| 0.25 | | X | X | X | X | X | X | | | |
| 0.50 | | X | X | X | X | X | X | | | |
| 1.00 | | X | X | X | X | X | X | | | |
| 1.50 | | X | X | X | X | X | X | | | |
| 2.50 | | X | X | X | X | X | X | | | |

5. DETERMINING OPTIMUM SLIT WIDTH AND THE EFFECT OF SLIT WIDTH ON MOTOR NOMINAL TORQUE

In this part, analyses were carried out to determine optimum slit width. The slit widths mentioned above were applied on the motor models having different slit depth and it was questioned which optimum slit width provides better motor torque. The obtained torque values for the motor models with 15 mm slit depth and different slit widths were given in Table 3. At the end of the analyses it was determined that the saturations occurring in the teeth increased as the slit width increased. It was determined that when the slit width value was especially at 0.25 mm and over, the saturations increased high enough to effect the torque. This was clearly shown with the torque values in Table 3.

Table 3. The obtained nominal torque and relative difference values for the motor models with 15 mm slit depth and different slit widths.

| Motor Models [mm] | Nominal Torque [N.m] | Relative Difference [%] |
|-------------------|----------------------|-------------------------|
| R.M. | 10.223 | |
| 0.09 | 10.241 | +0.176 |
| 0.10 | 10.632 | +4.000 |
| 0.15 | 10.555 | +3.247 |
| 0.25 | 10.055 | -1.643 |
| 0.50 | 10.262 | +0.381 |
| 1.00 | 9.597 | -6.123 |
| 1.50 | 9.372 | -8.324 |
| 2.50 | 7.317 | -28.426 |

The obtained torque-speed curves for the motor models with 15 mm slit depth and different slit widths are presented in Figure 8.

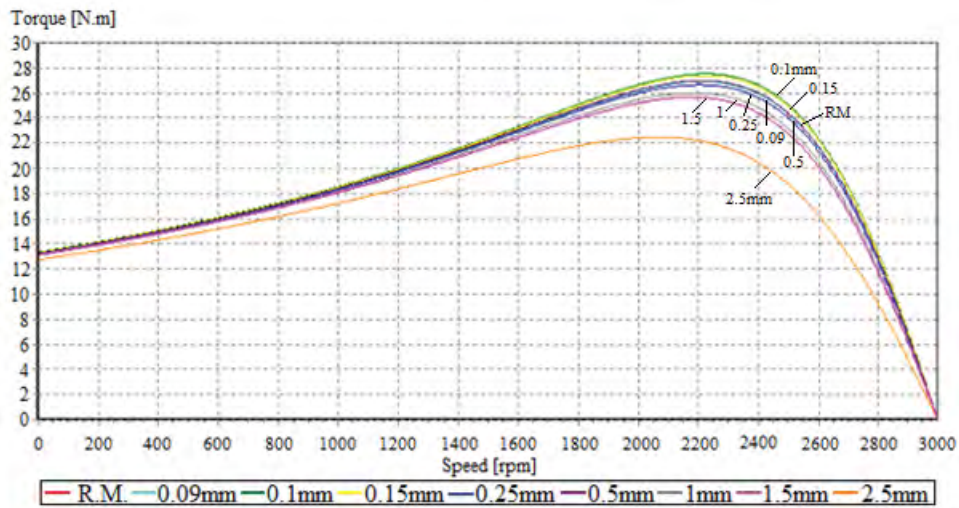


Figure 8. Torque - speed curves for the motor models with 15 mm slit depth and different slit widths.

When the table and figure were examined, for nominal torque, it was seen that the values obtained from 0.1 mm and 0.15 mm motor models were higher than the ones obtained from the other motor models and the reference motor model. At the end of the analyses, it was seen that the nominal torque value obtained from the motor model that had 0.09 mm slit width gave worse results than the value obtained from the 0.1 mm slit width which gave the best nominal torque value. It was seen that the nominal torque value decreased as the slit width increased depending on the increase of magnetic flow density amount. With these findings, it was determined that when the optimum slit width was 0.1 mm, motor nominal torque was at maximum level.

6. DETERMINING OPTIMUM SLIT DEPTH AND THE EFFECT OF SLIT DEPTH ON MOTOR NOMINAL TORQUE

Analyses were carried out at the following depths in order to determine optimum slit depth for each slit width. Analyses were carried out by creating models having from 6.25 mm to 23 mm slit depths in order to determine optimum slit depth. The results for different slit depths having 0.1 mm slit width that gave maximum value in this study were presented in tables and figures. The relative values according to reference motor models and nominal torque for 0.1 mm slit width different slit depths were given in Table 4.

Table 4. The obtained nominal torque and relative difference values for the motor models with 0.1 mm slit width and different slit depths.

| Motor Models [mm] | Nominal Torque [N.m] | Relative Difference [%] |
|-------------------|----------------------|-------------------------|
| R.M. | 10.223 | |
| 6.25 | 10.271 | +0.469 |
| 7.50 | 10.617 | +3.854 |
| 11.25 | 10.618 | +3.860 |
| 13.12 | 10.612 | +3.800 |
| 15.00 | 10.632 | +4.000 |
| 16.87 | 10.569 | +3.380 |
| 18.75 | 10.538 | +3.081 |
| 19.60 | 10.080 | -1.398 |
| 21.00 | 9.976 | -2.416 |
| 23.00 | 9.755 | -4.577 |

The obtained torque-speed curves for the motor models with 0.1 mm slit width and different slit depths are presented in Figure 9.

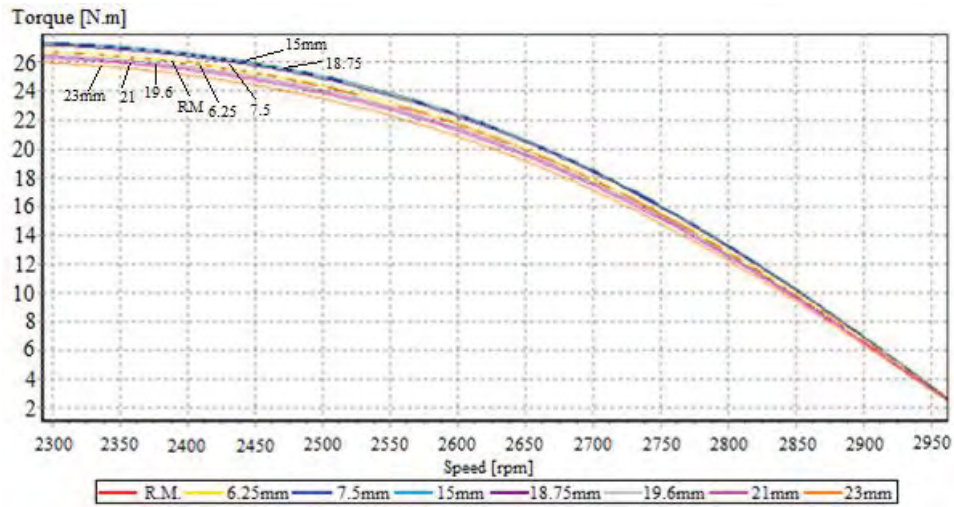


Figure 9. Torque - speed curves for the motor models with 0.1 mm slit width and different slit depths.

When the obtained results were examined, it was determined that minimum slit depth should be 7.5 mm and maximum slit depth should be 18.75 mm. It was seen that the motor models that had 7.5 mm, 11.25 mm, 13.125 mm, 15 mm and 18.75 mm slit depths had better nominal torque than other motor models. But the motor model that had 15 mm slit depth gave the best nominal torque value.

When the stator and rotor slot height values were considered, it was determined that the suitable slit value for both stator and rotor should be nearly the same height value as slot height value. It was observed that flux lines caused zigzag flux and completed their circuits around the slits in the models in which slit depth was smaller the slot height values. However, in the models in which the slit depth is higher, since the slits entered into the stator and rotor yokes, they caused saturation in these parts.

The nominal torque findings about the 48 different motor models produced for the slit optimization in the proposed slitted motor were presented in Table 5 and Figure 10. These findings easily showed the best value of the parameter researched.

Table 5. The nominal torque values obtained in different slit depth and width.

| Width of Slit [mm] | Depth of Slit [mm] | | | | | |
|--------------------|-----------------------------|--------|--------|---------------|--------|--------|
| | 7.50 | 11.25 | 13.12 | 15.00 | 16.87 | 18.75 |
| | Nominal Torque [N.m] | | | | | |
| 0.09 | 10.115 | 10.058 | 10.217 | 10.241 | 10.208 | 10.166 |
| 0.10 | 10.617 | 10.618 | 10.612 | 10.632 | 10.569 | 10.538 |
| 0.15 | 10.588 | 10.581 | 10.575 | 10.555 | 10.531 | 10.504 |
| 0.25 | 10.487 | 10.483 | 10.485 | 10.055 | 10.434 | 10.404 |
| 0.50 | 10.274 | 10.254 | 9.996 | 10.262 | 10.198 | 10.299 |
| 1.00 | 9.635 | 9.625 | 9.617 | 9.597 | 9.543 | 9.457 |
| 1.50 | 9.430 | 9.403 | 9.707 | 9.372 | 9.327 | 9.221 |
| 2.50 | 7.395 | 7.358 | 7.337 | 7.317 | 7.381 | 7.201 |

The nominal torque graphics obtained from slitted structure motor models were presented in Figure 11 according to their slit depth and width.

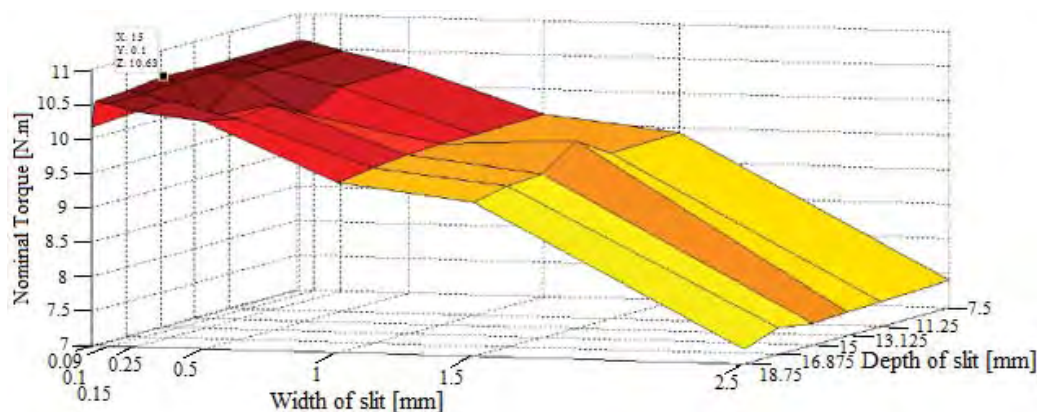


Figure 10. The nominal torque graphics for the motor models in different slit depth and width.

When the three dimensional graphics obtained from 48 different models were examined, it was seen that the motor model that had 15 mm slit depth and 0.1 mm slit width had the best nominal torque and as the slit depth and width increased the nominal torque value decreased.

7. CONCLUSIONS

In this study, a new teeth design is suggested so that the torque values which are one of the most important performance indicators of an induction motor could be improved. The re-engineered motor models with the suggested geometry are analyzed with FEMM software package and the following results were obtained.

It was found at the end of the analyses that when the optimum slit width value was 0.1 mm, motor nominal torque gave better results compared with other slit width values. The increase in slit width caused saturations in motor teeth and thus caused to worsen the motor torque - speed characteristic. It was decided at the end of different modeling that the depth of the slits used in the proposed slitted motor models should be almost the same height as stator and rotor height. Optimum slit value was determined as 15 mm. It was observed that flux lines completed their circuits around the slits in the models in which slit depth was smaller the slot height values and it negatively affected the motor torque. However, in the models in which the slit depth is higher, since the slits entered into the stator and rotor yokes, they caused saturation in these parts.

At the end of the modeling, 4% improvement was obtained in the motor models that had 15 mm slit depth and 0.1 mm slit width compared with the nominal torque values of reference model at nominal operation point.

Acknowledgement: This work was supported from Sakarya University Scientific Research Project.

8. REFERENCES

- [1] N. N. Soe, T. T. H. Yee and S. S. Aung, "Dynamic modeling and simulation of three phase small power induction motor," *World Acad Sci. Eng. Technol.*, vol. 42, pp. 421-424, 2008.
- [2] A. G. Yetgin, "Performance improvement of induction motor with slitted core design," Ph. D. Dissertation, Dept. Electr. and Electron. Eng., Sakarya Univ., Sakarya, Turkey, 2010.
- [3] M. Sundaram and P. Navaneethan, "On the influence of stator slot shape on the energy conservation associated with the submersible induction motors," *American J. App. Sci.*, vol. 8, no. 4, pp. 393-399, 2011.
- [4] M. Cunkas, "Intelligent design of induction motors by multiobjective fuzzy genetic algorithm," *J. Intell. Manuf.*, vol. 21, pp. 393-402, 2010.
- [5] F. Kentli, "A survey on design optimization studies of induction motors during the last decade," *Istanbul Univ. J. Elect. Electron. Eng.*, vol. 9, no. 2, pp. 969-975, 2009.

- [6] S. Manoharan, N. Devarajan, S. M. Deivasahayam and G. Ranganathan, "Review on efficiency improvement in squirrel cage induction motor by using DCR technology," *J. Elect. Eng.*, vol. 60, no. 4, pp. 227–236, 2009.
- [7] T. Aho, J. Nerg and J. Pyrhonen, "The effect of the number of rotor slits on the performance characteristics of medium-speed solid rotor induction motor," in *Power Electronics, Machines and Drives the 3rd IET Int. Conf.*, Dublin, Ireland, 2006, pp. 515-519.
- [8] J. Pyrhonen, J. Nerg, A. Mikkola, J. Sopanen and T. Aho, "Electromagnetic and mechanical design aspects of a high-speed solid-rotor induction machine with no separate copper electric circuit in the megawatt range," *Elect. Eng.*, vol. 91, pp. 35–49, 2009.
- [9] L. Li, A. Foggia, A. K. Lebouc, J. C. Mipo and L. Kobylansky, "Some armature reaction compensation methods numerical design of experiments and optimization for a hybrid excitation machine," in *Int. Electric Machines and Drives Conf.*, Miami, United States, 2009, pp. 832-838.
- [10] S. Chan and M. N. Hamid, "Finite element study on a two-phase switched reluctance motor with split rotor poles," in *Int. Conf. on Power Electronics and Drives Systems*, Malaysia, 2005, pp. 1156-1160.
- [11] W. L. Soong and N. Ertugrul, "Field-weakening performance of interior permanent-magnet motors," *IEEE Trans. Ind. Appl.*, vol. 38, no. 5, pp. 1251-1258, 2002.
- [12] A. G. Yetgin and M. Turan, "Efficiency optimization of slitted-core induction motor," *J. Elect. Eng.*, vol. 65, no. 1, pp. 60–64, 2014.
- [13] A. G. Yetgin, M. Turan and A. I. Canakoglu, "A novel slitted tooth core design to decrease leakage flux in induction motor," *J. Fac. Eng. Archit. Gazi Univ.*, vol. 27, no. 3, pp. 607-614, 2012.
- [14] T. Aho, J. Nerg and J. Pyrhonen, "Influence of rotor slit depth on the performance of the solid rotor induction motor," in *Energy Efficiency in Motor Driven Systems Conf. Proc.*, 2005, pp. 81–89.
- [15] T. Matsuo and T. A. Lipo, "Rotor dizayn optimization of synchronous reluctance machine," *IEEE Trans. Energy Convers.*, vol 9, no. 2, pp. 359-365, 1994.
- [16] M. E. Zaim, "Application of A nonlinear complex finite element method to the design of solid rotor reluctance machines," *IEEE Trans. Magn.*, vol. 34, no. 5, pp. 3592–3595, 1998.
- [17] M. Nashiki, A. Satake, Y. Kawai, T. Yokochi and S. Okuma, "A new flux-barrier-type reluctance motor with a slit rotor," *IEEE Trans. Ind. Electron.*, vol. 46, no. 6, pp. 1199–1206, 1999.
- [18] M. E. Zaim, "Non-linear models for the design of solid rotor induction machines," *IEEE Trans. Magn.*, vol. 35, no. 3, 1310-1313, 1999.
- [19] K. R. Dorairaj and M. R. Krishnamurthy, "Polyphase induction machine with a slitted ferromagnetic rotor: I-experimental investigations and a novel slipmeter," *IEEE Trans. Power Apparatus Syst.*, vol. 86, no. 7, pp. 835-843, 1967.
- [20] M. Jagiela and T. Garbiec, "Determination of best rotor length in solid-rotor induction motor with axial slitting," *Arch. Elect. Eng.*, vol. 61, no. 2, pp. 267-276, 2012.
- [21] Z. Zakaria, M. S. B. Mansor, A. H. Jahidin, M. S. Z. Azlan and R. A. Rahim, "Simulation of magnetic flux leakage (MFL) analysis using femm software," in *IEEE Symp. on Industrial Electronics and Applications*, Penang, Malaysia, 2010, pp. 481-486.
- [22] B.R. Singla, S. Marwaha and A. Marwah, "Design and transient analysis of cage induction motor using finite element methods," in *Int. Conf. on Power Electronics, Drives and Energy Systems*, New Delphi, 2006, pp. 1-5.
- [23] T. Lubin, K. Berger and A. Rezzoug, "Inductance and force calculation for axisymmetric coil systems including an iron core of finite length," *Prog. Electromagnet. Res. B*, vol. 41, pp. 377-396, 2012.

- [24] H. I. Lee, S. Y. Yoo and M. D. Noh, "Toroidally-wound self-bearing BLDC motor with lorentz force," *IEEE Trans. Magn.*, vol. 46, no. 6, pp. 2148-2151, 2010.
- [25] K. Boughrara, B. L. Chikouche, R. Ibtouen, D. Zarko and O. Touhami, "Analytical model of slotted air-gap surface mounted permanent-magnet synchronous motor with magnet bars magnetized in the shifting direction," *IEEE Trans. Magn.*, vol 45, no. 2, pp. 747-758, 2009.
- [26] N. Grilo, D. M. Sousa and A. Roque, "AC motors for application in a commercial electric vehicle: designing aspects," in *16th IEEE Mediterranean Electrotechnical Conf.*, Yasmine, Hammamet, 2012, pp. 277-280.
- [27] E. Saraiva, M. L. R. Chaves and J. R. Camacho, "Three-phase transformer representation using FEMM, and a methodology for air gap calculation," in *Proc. of the Int. Conf. on Electrical Machines*, Vilamoura, 2008, pp. 1-6.
- [28] D. Meeker, "Induction Motor Example," 2002.
- [29] A. İ. Çanakoğlu, İ. Şenol and D. N. Bekiroğlu, "Analytical and numerical calculation of inductance of a plunger-type magnet," in *Int. Conf. on Electrical and Electronics Engineering*, Bursa, Turkey, 1999, pp. 455-458.
- [30] A. G. Yetgin and N. Akbilek, "An analysis of stator and rotor leakage reactance in induction motors based on slot dimensions," in *11th Int. Conf. on Electrical Machines, Drives and Power Systems*, Sofia, Bulgaria, 15-16 September 2005, pp. 237-242.

Underwater Visual Tracking and Counting of Fishes

Md. Haidar Sharif¹, Sahin Uyaver², Adil Güler³

¹Gediz University, Izmir, Turkey

²Commerce University, Istanbul, Turkey

³Marmara University, Istanbul, Turkey

Abstract. We have addressed a method to use of computer vision techniques for underwater visual tracking and counting of fishes. The proposed method is a hybrid of background subtraction, Hungarian algorithm, and Kalman filter. It enables tracking of objects whose number may vary over time. Experimental results showed its effectiveness.

1 Introduction

Ecological observation is imperative for marine scientists to study marine ecosystems. Based on fish tracking, marine biologists are able to observe fish and their ecological environs. Normally, directly observing and quantifying fish behavior within a turbulence or trawl environment is challenging. In a traditional manner, marine biologists square off the existence and quantities of various types of fish using sundry methods e.g., casting nets in the ocean for collecting and examining fish, human underwater observation and photography[1], and combined net casting and acoustic (SONAR:SOund Navigation And Ranging)[2]. Fish detection and tracking are complicated by the variability of the underwater environment. A school of fishes can vary because fishes can enter a scene simultaneously and disappear. They can split or merge, as well as different school of fishes can be relatively near to each other or far away from each other. Because there exists no leader and every fish can know about local situation only around itself, the school of fishes should have a kind of intelligent ability as a whole[3]. Water plants are regarded as foreground objects as a result of the severe drift from the interference of the water flow, which results in complexities and difficulties in discriminating moving fish from drifting water plants. So the accuracy of fish tracking has been seriously affected if traditional methods are applied[4].

In recent years, considerable research has been conducted on video monitoring systems. Morais et al.[5] proposed a method, for underwater visual tracking and counting of fishes in vivo, based on the application of a Bayesian filtering technique that enabled tracking of objects whose number may vary over time. They demonstrated that their method could operate reliably under severe environmental changes (e.g. variations in water turbidity) and handle problems such as occlusions or large inter-frame motions. Wu et al.[6] proposed an algorithm to track fish group. The algorithm is based on the background subtraction

and Delaunay triangulation. An improved background subtraction algorithm was used to segment each frame of the video and then the center of each fish was evaluated. An optimized Delaunay triangulation network was established and the center of the fish group was calculated. Experimental results showed their algorithm could track fish groups accurately and provide effective data for monitoring water quality. Shao et al.[7] presented a real-time tracking system for moving objects on a water surface. Their tracking system was used to construct a fish-like robot water polo competition platform. The robot water polo game provided a real-time multi-agent environment from the viewpoint of distributed artificial intelligence and multi-robot cooperation research within water circumstance. A robust tracking algorithm based on template-matching method was proposed to overcome the specular reflections. Experimental results illustrated the effectiveness of their tracking algorithm. Shijun et al.[8] studied the impact of water quality on fish activity by tracking and measuring the fish swim speed. To achieve the tracking and test the velocity to the fish movement under the complex background, they proposed a method of video tracking algorithm consists of background subtraction and adaptive Kalman filter. They took special care to improve tracking accuracy and reduce loss of targets. Simulation results showed that the tracking accuracy of their proposed method. Tracking fish using implanted radio transmitters is an important part of studying and preserving native fish species. Jensen et al.[9] focused on developing simple methods for multi unmanned aerial vehicles navigation and transmitter localization. Swarm-like navigation methods (using potential fields) were used for such navigation, and a simple Kalman filter was used to estimate the location of the transmitter.

We have put forward a computer vision based algorithm to estimate and analyze fish trajectories and counting in high turbulence conditions. Our proposed method is similar to the proposed method of Wu et al.[6]. Our method used background subtraction, estimation of fish centers, Hungarian method, and Kalman filter, whereas Wu et al.[6] used improved background subtraction, estimation of fish centers, and Delaunay triangulation network. We have used the frequently used approach for discriminating moving target from background scene which is background subtraction. Background subtraction is a technique typically used to segment moving regions in image sequences taken from a static camera by comparing each new frame to a model of the scene background. The estimation of fish centers technique is similar to Wu et al.[6]. Then Hungarian algorithm has been used for data association. If we have m targets in any frame f_m and n targets in frame f_{m+1} , then Hungarian algorithm answers the question of data association: How do m targets relate to n targets? For example, if there are two targets $\{m_1, m_2\}$ in any frame f_m as well as two targets $\{n_1, n_2\}$ in the next frame f_{m+1} . Based on probability the Hungarian algorithm figures out the most accurate assignments from the set of assignments $\{m_1 = n_1, m_2 = n_2, \dots, m_1 \neq n_1, m_2 \neq n_2\}$. Finally, an adaptive Kalman filtering has been employed to observe data for self-testing when in the filtering, immutably judging the dynamic changes of the fish, and processing adaptive updates by observing data. Consequently, using videos obtained from fixed cam-

era the proposed method is able to estimate and analyze fish trajectories and counting them in high turbulence conditions over time.

The remaining part of this paper has been organized as follows: Section 2 delineates the proposed framework; Section 3 reports the experimental results; finally, Section 4 presents the conclusion of the work with few inklings for further investigation.

2 Proposed Approach

Flowchart of our proposed approach has been depicted in Fig. 1. Our proposed approach primarily includes five steps: (i) segmentation; (ii) fish center estimation; (iii) Hungarian algorithm for data association; (iv) tracking centers of fishes using Kalman filter; (v) get trajectories of fishes and counting them.

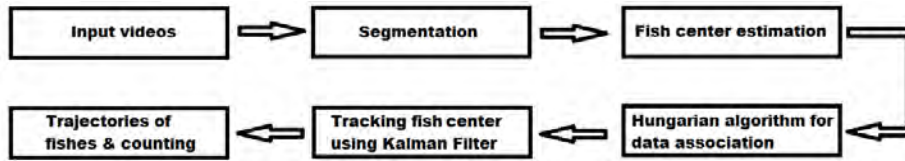


Fig. 1. Flowchart of our proposed approach.

2.1 Segmentation

Detecting moving objects is an important part in analyzing the scene. The most commonly applicable assumption is that the images of the scene with static objects exhibit some sort of regular behavior that can be well interpreted by a statistical model. If we have a statistical model of the scene, a moving object can be detected by spotting the parts of the image that do not fit the model. This technique is generally referred to as background subtraction. It is an old technique for finding moving objects in a video sequence. The scene model has a probability density function for each pixel separately. For a static scene the simplest model would be just an image of the scene without any moving object. A pixel from a new image is considered to be a moving pixel if its new value is not well described by its density function. Pixel values often have complex distributions and more elaborate models are required. Gaussian mixture model was proposed for background subtraction in [10]. One of the most commonly used approaches for updating Gaussian mixture model is presented in [11] and further elaborated in [12]. In this paper, the algorithm of [11] has been adopted.

2.2 Fish center estimation

On segmentation the contour of each fish has been detected. The center and the area of each contour can be estimated by the boundary points as performed by Wu et al.[6]. Let F be a fish centered at (x_0, y_0) and its size ω can be estimated by counting the pixels of F . To reduce the noise disturbance, dual threshold method has been used as [6] where T_1 and T_2 are minimum and maximum thresholds, respectively:

$$\omega = \begin{cases} 0, & \omega < T_1; \\ \omega, & T_1 \leq \omega \leq T_2; \\ 0, & \omega > T_2. \end{cases}$$

Let $\{(x_i, y_i) | i = 1, 2, \dots, n\}$ be a set of the boundary points of F then (x_0, y_0) can be estimated as: $x_0 = \sum_{i=1}^n x_i$ and $y_0 = \sum_{i=1}^n y_i$.

To track multiple fishes simultaneously for a long time, we have applied a Kalman filter approach based on center detection results of the fishes. Depending on the center detection results of the fishes, the number of detected fish blobs, obtained from segmentation, may not equal to the number of the initialized fishes. If the number of detected fish blobs is greater than that of existing fishes, then there may be some false detection or some new fishes. We can use some methods to initialize new Kalman filters to track the newly arrived fishes. Our method would work to track unfixed number of targets.

To decide which detection should guide which tracker, we need to solve a data association problem by assigning at most one detection to at most one fish. The optimal single-frame assignment can be obtained by using Hungarian Algorithm (sometimes referred to as the Munkres algorithm). A description of Hungarian Algorithm can be found in [13].

2.3 Data association and Kalman filter

The Hungarian method is a combinatorial optimization algorithm that solves the assignment problem in polynomial time and which anticipated later primal-dual methods. We have computed optimal assignment by Hungarian algorithm or so-called Munkres algorithm. The algorithm computes the optimal assignment (minimum overall costs) for the given rectangular distance or cost matrix. The result is a column vector containing the assigned column number in each row or 0 if no assignment could be done. The distance matrix may contain infinite values (forbidden assignments). Internally, the infinite values are set to a very large finite number, so that the Munkres algorithm itself works on finite-number matrices. Before returning the assignment, all assignments with infinite distance are deleted (i.e. set to zero).

To explain the Hungarian Algorithm a bit more clearly, we denote the problem as a bipartite graph. Fig. 2 gives an example of the bipartite graph when the number of detected fish blobs is larger than the number of the target being tracked. We have a complete bipartite graph $G = (V_d, V_f, E)$ with n detection vertices V_d at time t and m fish vertices V_f , and each edge E has non-negative

confidence $C(i, j)$. The bold blue lines in 2 illustrate the perfect matches with maximum confidence. In Fig. 2, $V_{d,t,i}$ are the detected blobs at time t , where $i = 1, 2, 3, \dots, n$, some of which may be false detection or newly arrived fishes. If a detection appears consecutively in some frames at almost the same position, and no existing tracker keeps tracking it, we can assume that a new fish appeared, and initialize a new Kalman filter tracker to track it. If the number of detected blobs is less than the previous number of fishes, some occlusion or fish disappeared might have occurred.

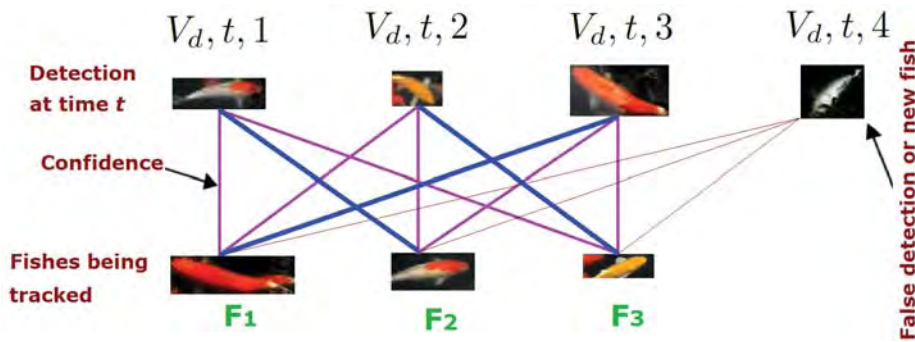


Fig. 2. Association between detected fish blobs and fishes being tracked.

3 Experimental Results

3.1 Data set

With a view to conducting experiments we have used fish related free videos from www.shutterstock.com.

3.2 Findings

A school of fishes tracking results have been depicted on the Fig. 3. Though there exists a very turbulence condition and fishes are varying over time, most of the fishes have been detected and tracked by the proposed algorithm. From the obtained trajectories, we have counted the number of fishes that crossed a defined region.

3.3 Shortcomings

From Fig. 3 it is clear that under a high turbulence condition over time a school of fishes are varying over time because some fishes are entering into the scene simultaneously and disappearing. Some of them are splitting or merging. As

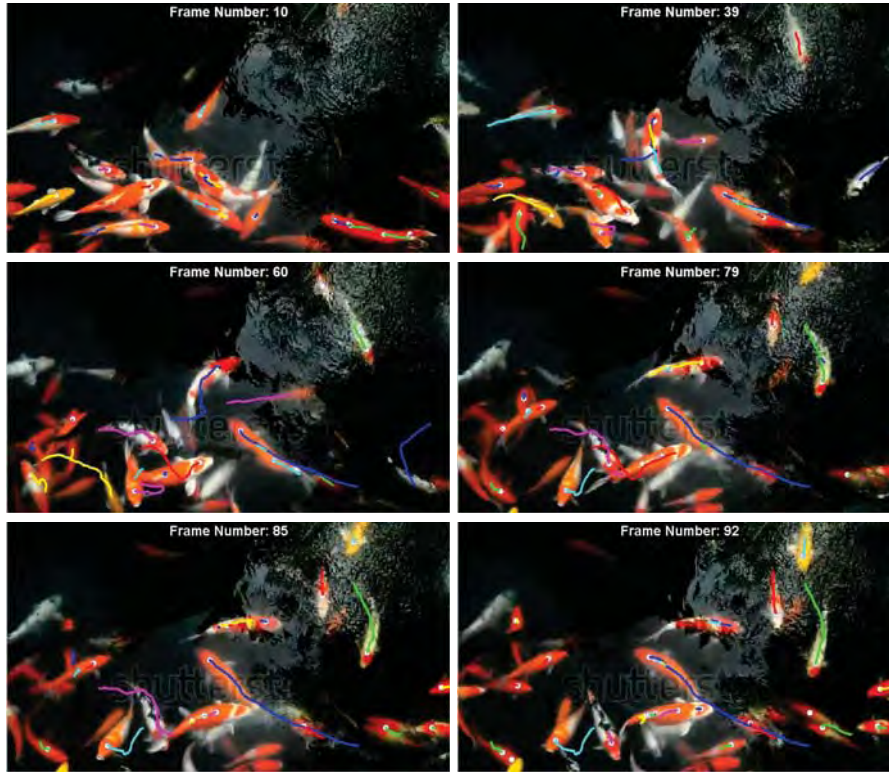


Fig. 3. Fish tracking results of the proposed algorithm.

a result, the algorithmic tracking and counting results have been affected in some extent. Future work would primarily focus to minimize or overcome this shortcoming of the proposed algorithm.

4 Conclusion

In this paper a simple and effective machine vision system capable of analysing underwater videos for detecting, tracking, and counting school of fishes has been presented. The method is based on background subtraction, Hungarian algorithm, and Kalman filter. It enables tracking of objects whose number would vary over time. Experimental results reported its efficacy.

References

1. Schlieper, C.: Research methods in marine biology. University of Washington Press, Seattle, 1972 [1](#)

2. Brehmera, P., Do Chib, T., D., M.: Amphidromous fish school migration revealed by combining fixed sonar monitoring (horizontal beaming) with fishing data. *Journal of Experimental Marine Biology and Ecology* **334**(1) (2006) 139–150 [1](#)
3. Hattori, K., Narita, Y., Kashimori, Y., Kambara, T.: Self-organized critical behavior of fish school and emergence of group intelligence. In: *International Conference on Neural Information Processing (ICONIP)*. Volume 2. (1999) 465–470 [1](#)
4. Shiau, Y., Chen, C., Lin, S.: Using bounding-surrounding boxes method for fish tracking in real world underwater observation. *International Journal of Advanced Robotic Systems* **10** (2013) 298 [1](#)
5. Morais, E., Campos, M.F.M., Padua, F., Carceroni, R.: Particle filter-based predictive tracking for robust fish counting. In: *Brazilian Symposium on Computer Graphics and Image Processing*. (2005) 367–374 [1](#)
6. Wu, J., Xiao, G., Zhang, Y., Chen, J., Zhu, L., Li, A.: Fish group tracking based on delaunay triangulation network. In: *International Congress on Image and Signal Processing (CISP)*. (2011) 534–537 [1](#), [2](#), [4](#)
7. Shao, L., Xie, G.: Real-time tracking of moving objects on a water surface. In: *International Conference on Mechatronics and Automation (ICMA)*. (2012) 2114–2119 [2](#)
8. Shijun, H., Yao, H.: A study of fish velocity measurement base on video tracking. In: *International Conference on Computer Science and Network Technology (ICCSNT)*. (2012) 1898–1901 [2](#)
9. Jensen, A., Chen, Y.: Tracking tagged fish with swarming unmanned aerial vehicles using fractional order potential fields and kalman filtering. In: *International Conference on Unmanned Aircraft Systems (ICUAS)*. (2013) 1144–1149 [2](#)
10. Friedman, N., Russell, S.: Image segmentation in video sequences: A probabilistic approach. In: *Thirteenth conference on Uncertainty in artificial intelligence*. (1997) 175–181 [3](#)
11. Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(8) (2000) 747–757 [3](#)
12. Elgammal, A., Harwood, D., Davis, L.: Non-parametric model for background subtraction. In: *European Conference on Computer Vision (ECCV)*. (2000) 751–767 [3](#)
13. Kuhn, H.W.: The hungarian method for the assignment problem. *Naval Research Logistics Quarterly* **2** (1955) [4](#)

Estimation of Targets from Satellite Images

Amira Youssef¹, Sahin Uyaver², and Md. Haidar Sharif³

¹The City of Scientific Research and Tecnological Applications, Alexandria, Egypt

²Commerce University, Istanbul, Turkey

³Gediz University, Izmir, Turkey

Abstract. Satellite sensor data have proven useful to the scientific communities. In this paper a simple and effective computer vision approach capable of analysing targets from satellite images has been addressed. The approach is based on the analysis of gray scale images obtained from satellite sensor and then further analysis of their connected component. Experimental results reported efficacy of our approach.

1 Introduction

Satellite imagery consists of images of Earth or other planets collected by artificial satellites. The use of satellite imagery in everyday life is by no means a novelty. Satellite images can provide huge amounts of data that in principle could be processed and provide very useful information in many areas. Satellite images are one of the most powerful and important tools used by the meteorologists, militaries, geographers, etc. For examples, satellite pictures help scientists to count number of penguins in Antarctica; or an authority would be interested to count the number of buildings in some region.

Numerous works can be found in the literature that directly concern to the detection of targets from satellite images. For instances, Aytekin et al.[1] presented a bag of visual words algorithm for object detection in satellite images. In their algorithm steps a new descriptor was presented by adding scale information to SIFT descriptor and a novel visual word weighting algorithm was proposed considering that more occurrence in the object and less in the background was an importance measure. A new concept for the detection of small objects from modular optoelectronic multispectral scanner (MOMS-02) high spatial resolution panchromatic satellite imagery was presented by Segl et al.[2]. The authors combined supervised shape classification with unsupervised image segmentation in an iterative procedure which allowed a target-oriented search for specific object shapes. Pirzada et al.[3] compared edge detection based on bilateral filtering with canny edge detection technique for satellite images. Their proposed bilateral filtering based edge detection not only generated well localized edges but also simultaneously reduced considerable noise from real life images. Their results showed that the bilateral filtering based edge detection provide better edge maps than other comparable techniques. Ke et al.[4] proposed an automatic and rapid method to detect objects from satellite image with large size, which is the precondition for detailed object recognition. Their feature based method consists

of some Haar-like structural features with the help of Adaboost classifier. They claimed that the detected object by using their algorithm, object's details could be further recognized. However, pixel based method usually performs slower than feature based method, in many applications where execution time is not a big factor we can use pixel based method for better performance.

We have put forward a computer vision based algorithm to estimate the number of targets (e.g., buildings, trees, penguins, etc.) from satellite images. Our pixel based method deals with gray scale images and construct histogram thereof. Thus any kind of color satellite image will be converted to gray scale as well as obtained its complement and then a histogram will be built to get the knowledge of thresholds. On constructing histogram, the image will be divided into pieces based on a set of thresholds. A standard connected components analysis algorithm has been applied on it. The connected component algorithm works by looping over the pixels of the divided image. Any set of pixels which is not separated by a boundary is call connected. Each maximal region of connected pixels is called a connected component, which gives the estimation information of target.

Although the method is computationally expensive, it is able to estimate the number of targets from some given satellite images, which primarily include buildings or trees or penguins or so on.

The remaining part of this paper has been organized as follows: Section 2 delineates the proposed framework; Section 3 reports the experimental results; finally, Section 4 presents the conclusion of the work with few inklings for further investigation.

2 Proposed Approach

Flowchart of our proposed approach has been depicted in Fig. 1. Our proposed approach primarily includes five steps: (i) Get gray scale image from any kind of colored satellite image; (ii) A histogram is built to get the knowledge of thresholds; (iii) Apply a set of thresholds; (iv) Connected component analysis; (v) Target estimation.



Fig. 1. Flowchart of our proposed framework

2.1 Histogram from gray scale image

Fig. 2 shows the image and histogram for a satellite image. This image has been collected from internet. However, the histogram shows that the vast majority of the pixels are of medium intensity. Most of the building in this image is a shade of dark gray. There are several buildings with high intensity. It looks that a single threshold would not do a good job in isolating those buildings.

2.2 Thresholds

Threshold values should be in the intensity interval of 0 (black) to 255 (white). We use a minimum and a maximum thresholds to create a binary image that is between the pair of minimum and maximum intensity values and all other values map to 0 (black).

2.3 Connected component analysis

After constructing histograms, we have divided images into pieces based on a set of thresholds and applying a connected components analysis algorithm.

A standard connected component algorithm has been taken into account and it takes as input the binary image obtained by threshold proceeding as well as label image structure. The structure consists of a two dimensional array which is the label image, an array of label sizes, and the maximum label value. The array of sizes tells us how big (number of pixels) a certain component is, and is indexed by label number. The maximum label value has been used for calling flood fill and it also tells us how many components are currently in the image.

The connected component algorithm works by looping over the pixels of the binary image. If that pixel is 1 (true) in the binary image and -1 in the label image, flood fill is called with the appropriate data. If that pixel is 0 (false) in the binary image, the pixel is changed to zero in the label image (no component). Once connected component has been finished, the label image will have no -1 pixels. Every pixel is either 0 or a positive integer.

Since we are interested in dividing images into pieces, each of the algorithm must be run multiple times. Every label has a random color generated by a composition of red, green, and blue component between 0 and 255.

2.4 Target estimation

Any component that is smaller than a defined number of pixels (say 250) will be merged or removed. The rest of the components will be counted as the targets.

3 Experimental Results

With a view to conducting experiments we have used several satellite images collected from internet.

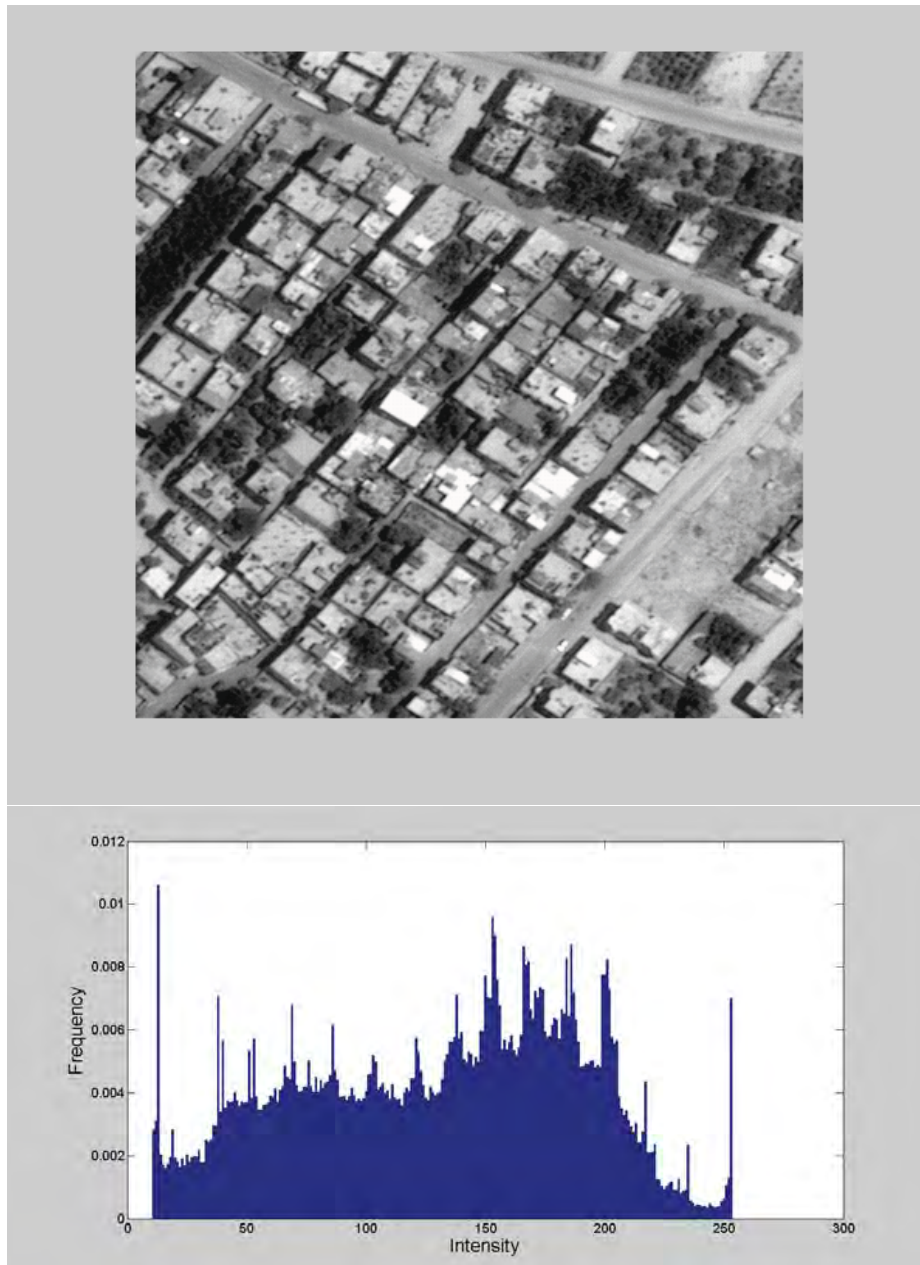


Fig. 2. A gray scale image and its corresponding histogram.

Fig. 3 depicts a satellite image. We are expecting to estimate the number of buildings of it. The algorithm has executed and provided the knowledge of estimation with varying parameters. For examples, first row of Fig. 3 with minimum 80 and maximum 110 thresholds gave the number of connected components or targets 7130. But it is very noisy. Thus we have applied a threshold of 250 pixels to filter the output. If any connected component which is smaller than 250 pixels then that component will be merged its nearest neighboring connected component or moved to background. Finally, the algorithm gave 66 connected components which are the estimated targets in that image. But this estimation is far from the ground truth. Consequently, we need to adjust the thresholds. Table 1 shows the detailed view of the detection results.

| Thresholds [min max] | Targets before filter | Filtered pixels | Targets after filter |
|----------------------|-----------------------|-----------------|----------------------|
| [80 110] | 7130 | 250 | 66 |
| [60 120] | 4010 | 240 | 113 |
| [40 150] | 2436 | 230 | 82 |
| [90 140] | 5452 | 250 | 124 |

Table 1. Threshold effect on the detection results

From Fig. 3 and its associated Table 1 it is clear that threshold set [60 120] with rejecting pixels threshold 240 gives us the best estimation of the targets in the image.

4 Conclusion

In this paper a simple and effective computer vision approach capable of analysing targets from satellite images has been presented. Though the method is computationally expensive, it is able to estimate the number of targets from some given color satellite images.

References

1. Aytakin, C., Alatan, A.: A novel bag of visual words model for object detection in satellite images. In: Signal Processing and Communications Applications (SIU), 2011 IEEE 19th Conference on. (April 2011) 654–657 [1](#)
2. Segl, K., Kaufmann, H.: Detection of small objects from high-resolution panchromatic satellite imagery based on supervised image segmentation. *Geoscience and Remote Sensing, IEEE Transactions on* **39**(9) (Sep 2001) 2080–2083 [1](#)
3. Pirzada, S., Siddiqui, A.: Analysis of edge detection algorithms for feature extraction in satellite images. In: Space Science and Communication (IconSpace), 2013 IEEE International Conference on. (July 2013) 238–242 [1](#)
4. Ke, Y., e.a.: A rapid object detection method for satellite image with large size. In: International Conference on Multimedia Information Networking and Security (MINES). (2009) 637–641 [1](#)

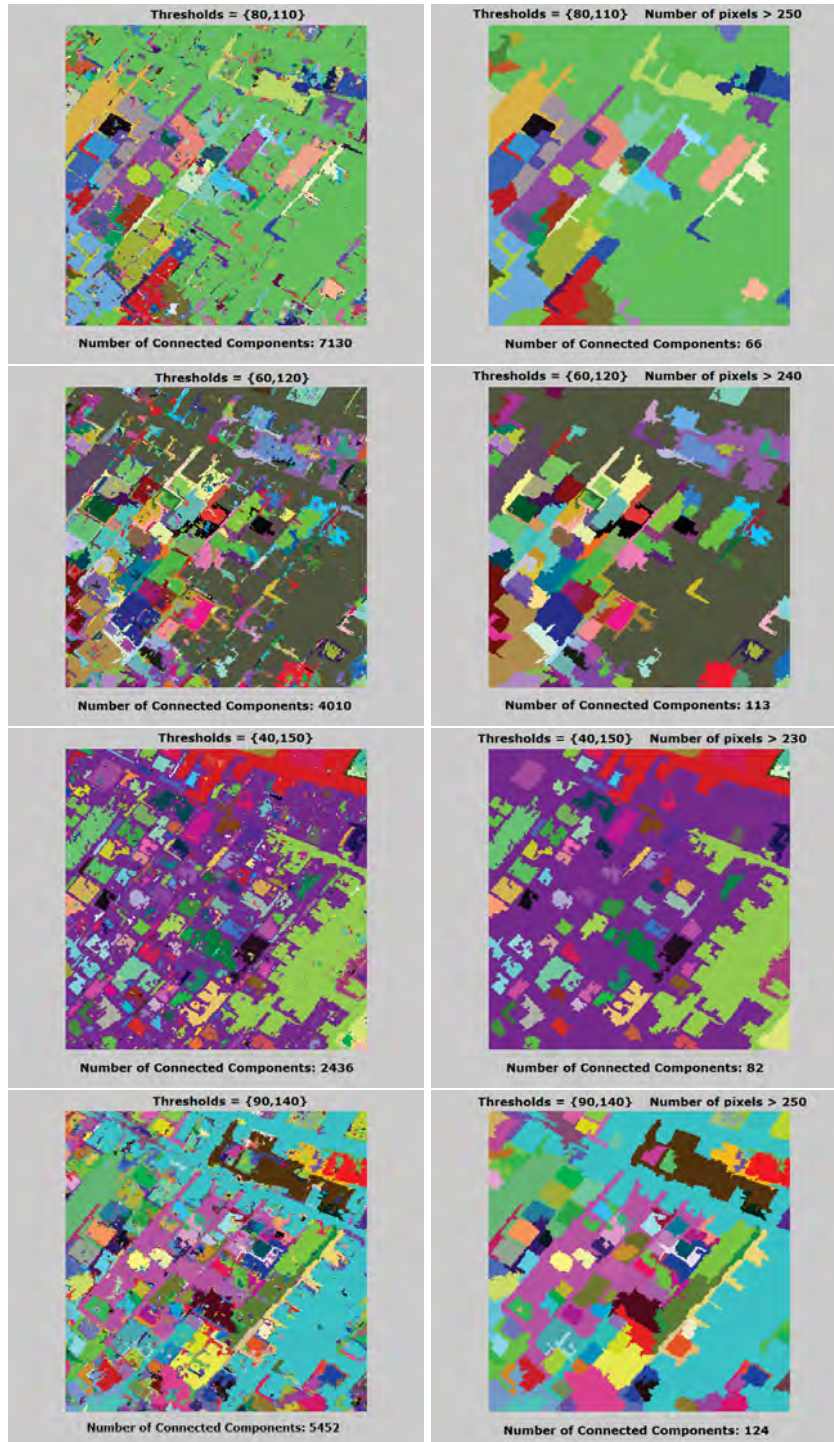


Fig. 3. Target estimation results of the proposed algorithm.

A Meta-Heuristic Approach for Optimal Train Control

Kemal Keskin¹, Abdurrahman Karamancioglu²

^{1,2}Electrical Engineering Department, Eskisehir Osmangazi University,
Eskisehir, TURKEY

¹kkeskin@ogu.edu.tr, ²akaraman@ogu.edu.tr,

ABSTRACT

Environmental concerns and growing demand for energy has increased significance of energy-efficient train operations. In this manuscript, an energy-efficient train operation based on optimal selection of switching times for control inputs is studied. One of the metaheuristic approaches, genetic algorithm (GA), is used to find the switching times for optimal train operation. In our problem formulation, the objective function comprises energy consumption, travel time and maximum speed with penalty factors. In order to verify the optimization results we obtained, a simulation is performed using MATLAB. For the simulations, two test tracks are created with different grade profiles and track lengths. Two strategies are simulated on each of these test tracks. The first strategy is the straightforward approach consisting of maximum acceleration and braking phases. The second strategy is combination of maximum acceleration, cruising, coasting and braking phases where optimal switching times are calculated by GA. Simulation results show that, compared to the first strategy, the second strategy can save energy up to 30% while travel time increases by 5%.

Keywords – train motion, energy-efficient control, genetic algorithm, grade profile, simulation of train

1. INTRODUCTION

Increasing demand for energy and environmental concerns makes energy-efficient transportation important for the world. Railway systems can satisfy high capacity travel demand with low energy consumption and without significant loss of service quality. However, railway systems requires vast amount of energy during operation. So it is possible to save remarkable amount of energy by making small changes in any part of the system or in its operation strategies. Milroy [4] proposed a controller for the minimization of energy consumption subject to timetable and operational constraints. In 1990, with the consideration of discrete control, Howlett [5] showed that optimal strategies can be found if the problem is defined in suitable function spaces. Khmelnsky [6] presented an analytical solution to optimal control problem based on maximum principle analysis. He demonstrated the solutions' accuracy with numerical examples. Besides these researches, Liu and Golovitcher [7] proposed an analytical solution which gives the sequence of optimal controls and equations to find the control change points. They developed an algorithm and a computer program for energy efficient train control. In order to implement energy saving programs and operation strategies, Jong and Chang [8] presented two models which estimates energy consumption of a train. They verified the proposed models in a real railway with electric train from Taiwan Railway Administration. In [9] authors used a measure theory technique and iterative dynamic programming algorithm to solve energy minimization problem. Chang and Sim [10] proposed a dynamic controller based on finding coasting points. Then they applied GA to determine when coasting should be initiated or terminated. In another study Wong and Ho [11] used GA to search for appropriate coasting points. Kim and Chien [12, 13] developed a simulated annealing algorithm to search for optimal train operation, considering track alignment and speed limits.

The structure of this paper is as follows: In Section 2, A train motion is modeled and consumed energy is defined in terms the motion variables. In Section 3, train operation strategies are defined

and energy optimization problem is introduced. In Section 4, a case study and its results are given. Last section consists of conclusions.

2. MODELING THE MOTION

A train can be modeled in different ways. Among them two main approaches which are widely used in the literature: Distributed mass model and single mass-point model [2]. In this research, we consider only the energy minimization, therefore, a mass-point train model is more convenient. We consider a train motion between two stations which constitutes train's basic motion in a journey. We let the distance between stations be X and allowed journey time be T . Applying Newton's second law to the train one may write the motion equation of train as

$$\frac{dx}{dt} = v(t) \quad (1)$$

$$\frac{dv}{dt} = \frac{F(v) - F_b(v)}{m} - R(v) - R_g(x) - R_c(x) \quad (2)$$

where x is the position and v is the velocity. They are constrained by $0 \leq x \leq X$, $0 \leq v \leq v_{max}$. The symbols m, F, F_b, R, R_g and R_c denote mass of the train, traction effort, brake effort, and resistances against to the train motion respectively. Resistance of train, R , can be calculated by using Davis equation [1].

$$R = A + Bv + Cv^2$$

In the equation above, the coefficients A and B correspond to mass and mechanical resistance and C corresponds to air resistance. These coefficients depend on train characteristics and external forces. Level changes in track causes gradient resistance, R_g , and can be calculated as

$$R_g = m \cdot g \cdot \sin \alpha$$

where g is gravitational constant and α is the gradient. The term R_c in equation (2) represents the resistance of curves. It depends on train characteristics and radius of the curve.

Traction effort provides force to move train along the rail line. It is important to obtain desired velocity and needs to be calculated in terms of acceleration rates. Traction effort is restricted to certain limits due to adhesion between wheel and rail surfaces. It is calculated by

$$F = 2650 \cdot \frac{\mu \cdot P}{v}$$

where μ is the efficiency in converting motor power to traction force, P is locomotive's horse power, v is the velocity of train and 2650 is for unit conversion [3]. By using this equation, traction effort and also the consumed power can be calculated for each time step. The consumed power, P is defined by the equation below.

$$P = \frac{F \cdot v}{2650 \cdot \mu}$$

to find total energy consumption, power is integrated over time.

$$E = \int_0^T P dt \quad (3)$$

In the next section we present a discrete model for optimizing the energy as a function of cruising and coasting positions.

3. TRAIN OPERATION AND ENERGY OPTIMIZATION

Energy-efficient operation of trains can be improved in many ways such as: by reducing resistances or increasing engine efficiency. Since these kind of approaches have physical limits and additional cost, switching driving scenarios is mostly used instead. Basically driving scenarios consists of the following elementary motions: maximum acceleration, cruising, coasting, and braking. Figure 1 shows a position versus velocity graph for an example scenario.

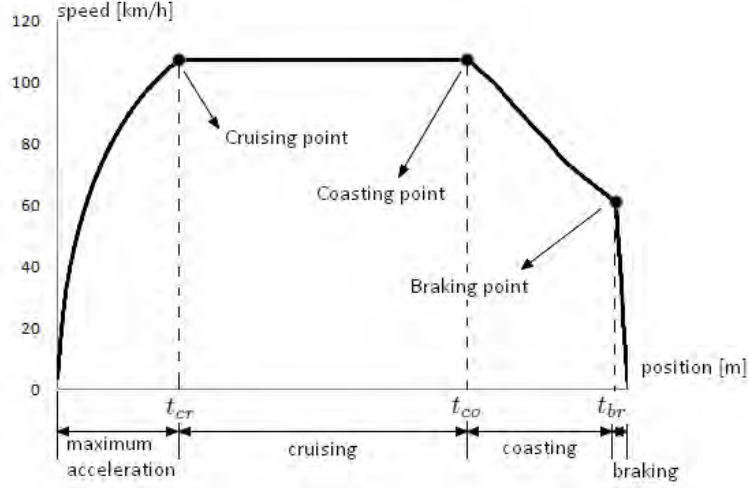


Figure 1. Motion phases: Maximum acceleration, cruising, coasting and braking

In Figure 1, maximum acceleration is applied to train at the beginning of travel. Notice that, at some point, cruising starts and train continues its travel with constant speed. Another phase, coasting means to move along under obtained momentum. For energy efficiency, coasting is applied between cruising and braking points. This phase persists until train reaches the stopping distance. *Stopping distance* is a function of remaining distance and current velocity of train. Braking should start at this point such that train can stop at station safely.

For the system dynamics in differential equations (1) and (2), state variables can be defined in the time interval $[0, T]$ for N steps as,

$$\left. \begin{aligned} v_{k+1} &= v_k + a_k \cdot \Delta t \\ x_{k+1} &= x_k + \left(\frac{v_k + v_{k+1}}{2} \right) \cdot \Delta t \end{aligned} \right\} k = 0, 1, \dots, N$$

where v_k , a_k , x_k represent velocity, acceleration, position of train at step k respectively. Δt denotes step size and equals T/N . For $k \in \{0, 1, \dots, N\}$, this dynamics determines energy at k -th step, and the total energy consumed throughout the trip is the sum of energy consumption at each step:

$$\sum_{k=1}^N \frac{F \cdot v_k}{2650 \cdot \mu} \cdot \Delta t \quad k = 0, 1, \dots, N \quad (4)$$

The total energy can be decomposed into its components for every phase. In the *maximum acceleration* phase, the traction effort is fixed at its attainable maximum value, therefore speed monotonically increases. This phase starts at $k = 0$ and ends at $k = k_1$, where optimal value of k_1 to be determined in the optimization process. In the second phase - *cruising phase* - the speed is fixed to a value, v_{k_1} , where F_k gets different values which are equal to resistance values. This phase extends between k_1 and k_2 , where optimal k_2 to be determined. For the subsequent phases (coasting and braking phases), no energy is consumed due to zero traction force. Defining $t_{cr} = k_1 \Delta t$ and $t_{co} = k_2 \Delta t$ we can write the objective function and the constraints of the optimization problem as follows:

$$\begin{aligned} & \min_{t_{cr}, t_{co}} \left(\sum_{k=0}^{k_1} \frac{F_{max} v_k}{2650 \mu} \Delta t + \sum_{k=k_1+1}^{k_2} \frac{F_k v_{k_1}}{2650 \mu} \Delta t \right) \\ & 0 \leq F_k \leq F_{max} \\ & 0 \leq v_k \leq v_{max} \quad v(0) = 0 \quad v(N) = 0 \\ & (N - k_2) \Delta t \leq \text{stopping distance} \\ & k_1 \leq k_2 \leq N \end{aligned}$$

In the next section we solve the above optimization problem for a case with various track conditions and compare the consumed energy results to that of the straightforward driving approach.

4. A CASE STUDY

The problem of energy-efficient train operation is studied for two test conditions presented in this section. Since not only it can handle with multiple local optima but also it is modular, for the computation of switching points, a genetic algorithm method is implemented on MATLAB. Using the calculated switching points, we simulated the train motion and showed the corresponding energy consumptions below.

Simulations and Results

The proposed approach is applied to two different test cases. In Test 1, there is a track with an uphill part followed by a downhill part, in Test 2, there is a track with an uphill part followed by a downhill part and a downhill part followed by an uphill part. Changes in level with regard to track length and also slope rates are given in Figure 2.

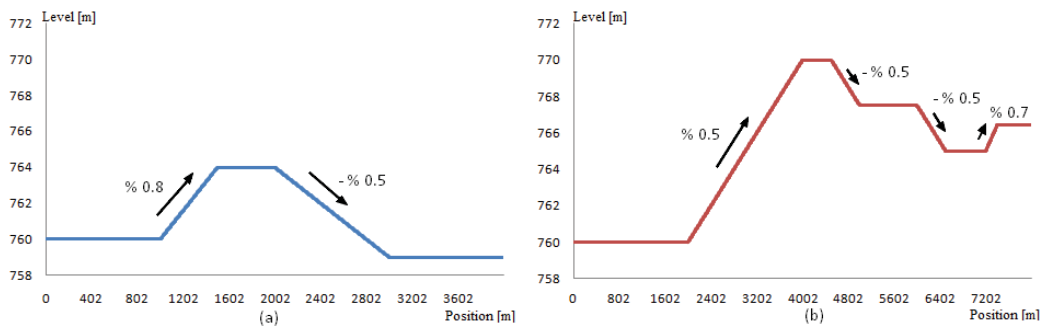


Figure 2. Track length versus track level (grade profiles) (a) for Test 1 (b) for Test 2

We used a MATLAB program to calculate and simulate speed profiles. In this program, a train considered has 3300 hp power and its maximum operating speed is 131 km/h. Also maximum acceleration/deceleration rate is restricted to $2 m/s^2$. The train is initially at zero position and has 0 km/h velocity.

Test 1

A simple 4000 m length track with an uphill and a downhill section is defined for this test. Two control strategies which depend on phases defined in Figure 1 are applied and results are compared. *Strategy 1* consists of only *maximum acceleration* and *braking* phases. This strategy corresponds to driving train in maximum acceleration mode until the safe stopping distance. *Strategy 2* is a combination of all phases mentioned before: *maximum acceleration* + *cruising* + *coasting* + *braking*. This strategy is applied in three different experiments: (a) when travel time is restricted to 158 s, (b) when travel time is restricted to 175 s, (c) when travel time is restricted to 200 s. Table 1 shows the obtained simulation results.

| TEST 1 | Strategy 1 | Strategy 2 (a) | Strategy 2 (b) | Strategy 2 (c) |
|--------------------------|------------|----------------|----------------|----------------|
| Duration (seconds) | 151 | 158 | 174 | 200 |
| Energy Consumption (kwh) | 102.40 | 70.42 | 55.72 | 45.06 |

Table 1. Simulation durations and energy consumptions for Test 1

Table 1 shows that, using the first strategy, the train reached the target distance in 151s and with 102.40 kwh energy consumption. For the second strategy, train reached the target distance: (a) in 158s and with 70.42 kwh energy consumption, (b) in 174s and with 55.72 kwh energy consumption, and (c) in 200s and with 45.06 kwh energy consumption. It is obviously seen from results that duration increases while energy consumption decreases. But also it is quite clear that while losing 5% time, energy can be saved up to 30%.

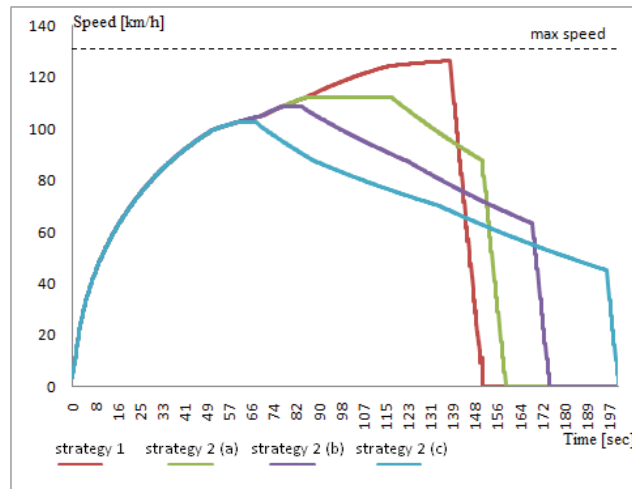


Figure 3. Optimal speed profiles versus time for Test 1

Test 2

An 8000 m length track with an uphill followed by a downhill section and a downhill followed by an uphill section is defined for Test 2. Two control strategies are applied as Test 1 and results are compared. *Strategy 2* is applied also for three different allowable times: (a) when travel time is restricted to 273 s, (b) when travel time is restricted to 282 s, (c) when travel time is restricted to 310 s. Test results are presented in Table 2.

| TEST 2 | Strategy 1 | Strategy 2 (a) | Strategy 2 (b) | Strategy 2 (c) |
|--------------------------|------------|----------------|----------------|----------------|
| Duration (seconds) | 266 | 273 | 282 | 310 |
| Energy Consumption (kwh) | 180.68 | 147.73 | 137.87 | 117.37 |

Table 2. Simulation durations and energy consumptions for Test 2

As in the first test the duration increases while energy consumption decreases. It is also seen that while losing 2.6% time, energy can be saved up to 18.3%.

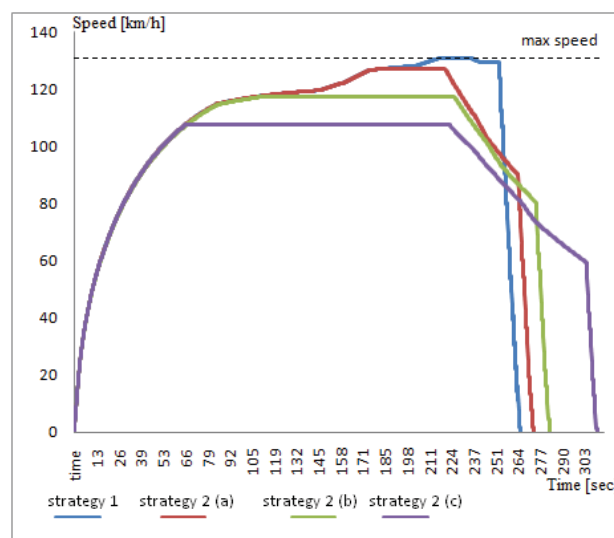


Figure 4. Optimal speed profiles versus time for Test 2

5. CONCLUSIONS

In this study, an optimization approach for an efficient operation of a train, based on metaheuristic method, is proposed. A genetic algorithm is implemented on MATLAB. The objective function is defined as energy consumption throughout the journey. The simulation results demonstrated that with appropriate speed profile energy consumption decreases significantly. The optimization model developed in this study can also be generalized to more general sets of constraints in a straightforward way.

5. REFERENCES

- [1] Davis, T., "Transportation Energy Data Book", 20th ed. Washington, DC, 2000
- [2] Howlett, P.G. and Pudney, P.J., "Energy-Efficient Train Control", Advances in Industrial Control, Springer, London, 1995
- [3] Hay, W.W., "Railroad Engineering", 2nd Ed., Wiley, New York, 1982.
- [4] Milroy, I.P., "Aspects of automatic train control", Ph.D. thesis, Loughborough University, Australia, July 1980
- [5] Howlett, P.G., "An optimal strategy for the control of a train", ANZIAM Journal, Vol. 31, 1990, pp 454-471
- [6] Khmelnitsky, E., "On an optimal control problem of train operation", *IEEE Transactions Automatic Control*, Vol. 45 Issue 7, 2000, pp 1257-1266
- [7] Liu, R. and Golovitcher, I., "Energy-efficient operation of rail vehicles", *Transportation Research Part A: Policy and Practice*, Vol. 37, 2003, pp 917-932
- [8] Jong, J. and Chang, E., "Models for estimating energy consumption of electric trains", *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, 2005, pp 278-291
- [9] S. E_ati, H. Roohparvar, "The minimization of the fuel costs in the train transportation", *Applied Mathematics and Computation*, Vol. 175 Issue 2, 2006, pp 1415-1431
- [10] C.S. Chang and S.S. Sim, "Optimising train movements through coast control using genetic algorithms", *Electric Power Applications*, IEE Proceedings, Vol. 144 Issue 1, 1997, pp 65-73
- [11] K. K. Wong, T. K. Ho, "Dynamic coast control of train movement with genetic algorithm", *International Journal of Systems Science*, Vol. 35, 2004, pp 835-846
- [12] K. Kim, S.I. Chien "Simulation-Based Analysis of Train Controls under Various Track Alignments", *Journal of Transportation Engineering*, Vol. 136, 2010, pp 937-948
- [13] K. Kim, S.I. Chien "Optimal Train Operation for Minimum Energy Consumption Considering Track Alignment, Speed Limit, and Schedule Adherence", *Journal of Transportation Engineering*, Vol. 137, 2010, pp 665-674

A Hybrid Algorithm for Exam Timetabling Problem in Marmara University

M. E. Gedikli¹, M. Agaoglu²

¹ Faculty of Engineering, Computer Engineering Department, Marmara University,
Kadikoy, Istanbul, TURKEY

² Faculty of Engineering, Computer Engineering Department, Marmara University,
Kadikoy, Istanbul, TURKEY

¹emin.gedikli@marmara.edu.tr, ²agaoglu@marmara.edu.tr

ABSTRACT

Exam timetabling problem, a well-studied combinatorial optimization problem, is one of the important administrative activities of universities and other educational institutions. There are a lot of different techniques to solve this problem. Some of these techniques use only one algorithm and others combine different algorithms. We solve the exam timetabling problem of Marmara University using a hybrid algorithm. Four local search based meta-heuristics, Tabu Search, Simulated Annealing, Hill Climbing and Great Deluge algorithms are used. Every algorithm has its own structure and parameters. It is necessary to fine-tune the basic parts of each algorithm for generating successful solutions in hybrid approaches. The application for the exam timetabling problem is implemented in Marmara University. It is observed that the hybrid algorithm gives more successful results.

Keywords – exam timetabling, great deluge, hill climbing, hybrid algorithm, simulated annealing, tabu search.

1. INTRODUCTION

Scheduling is the process of allocating resources to activities with considering constraints about the resources. Wren (1996) defines the objective of scheduling is to solve practical problems relating to allocation, subject to constraints, of resources to objects being placed in space-time, using or developing whatever tools may be appropriate.

Timetabling problems, can be thought of a type of scheduling problems, become also a wide research area that ranges from educational timetabling (Burke, de Werra, et al., 2004), nurse scheduling (Burke, De Causmaecker, et al., 2004), sports timetabling (Easton et al., 2004), employee timetabling to transportation timetabling (Kwan, 2004).

Burke, de Werra, et al. (2004) give a definition of general timetabling, which covers many cases:

“A timetabling problem is a problem with four parameters: T , a finite set of times; R , a finite set of resources; M , a finite set of meetings; and C , a finite set of constraints. The problem is to assign times and resources to the meetings so as to satisfy the constraints as far as possible.”

At this definition, T , R , M and C terms can be explained as below:

A *time t* is an element of the set of times T of an instance of the timetabling problem. A *time period* is a variable constrained to contain one time.

A *resource r* is an element of the set of resources R of an instance of the timetabling problem. A *resource slot* is a variable constrained to contain one resource.

A *meeting m* is a named collection of *time periods* and *resource slots*. Assigning values to these slots means that all of the assigned resources attend this meeting at all of the assigned times.

A *constraint* c is an element of the set of constraints C of an instance of the timetabling problem. Constraints include hard and soft constraints.

A timetabling problem can be defined to be the problem of assigning a number of events into a limited number of time periods (Burke & Petrovic, 2002).

Educational timetabling is one of the most studied parts of timetabling problems. It deals with assigning time periods, instructors, students and rooms to a set of meetings in such a way that no student and instructor has to attend two meetings simultaneously. In universities and other educational institutions, educational timetabling is a basic and major activity for each semester in every year.

Educational timetabling can be classified into two main headings: course timetabling and exam timetabling. In exam timetabling, exams are assigned to a number of available classes, rooms, halls, laboratories (of certain capacity) and a number of periods (timeslots) considering the constraints.

It can be thought that course and exam timetabling are same problems, but important differences exist between these. For example, a course usually has to be assigned into one room, while an exam may be split to more than one room. Also, a course may be split to more than one period at different days, but an exam usually has to be assigned to one period.

Also exam timetabling problem differs from the course timetabling problem at the following points (Schaerf, 1999):

- There is only one exam for each course per exam weeks.
- There can be more than one exam assigned to one room.
- The conflict condition for exam timetabling is generally stricter than course timetabling problem. It can be accepted that a student is forced to skip a lecture because of overlapping, but not that a student skips an exam.
- In exam timetabling, it is usually considered to spread the exams out as much as possible. On the other hand, for course timetabling it is often considered undesirable to spread the courses out. Students tend to prefer to have courses in contiguous blocks (Burke, de Werra, et al., 2004).
- There are different types of constraints, e.g. at most n exam per day for each student, and not too many consecutive exams for each student.
- The number p of periods may differ, in contrast to course timetabling where it is fixed.

2. CONSTRAINTS

In the timetabling literature, there are some constraints to be considered when solving timetabling problems. These constraints can be divided into two categories: *hard constraints* and *soft constraints*.

2.1 Hard Constraints

Hard Constraints are constraints that have to be satisfied under any circumstances. Solutions can't violate the hard constraints. These solutions are called "feasible solutions". For example, no student can be assigned to two exams at any one time period. In the literature, there are a lot of hard and soft constraints. Common hard constraints are as follows:

- The total resources that are required in each time period can't be greater than the resources that are available.
- No resource (student or instructor) can be enforced to be in more than one place at any one time.
- No exams with common resources (student) can be assigned simultaneously.
- Students have n exams at most one day.

2.2 Soft Constraints

Soft Constraints are constraints that need not to be strictly satisfied. They are wanted to satisfy, but not essential. Soft constraints have penalty costs that we should minimize. There may be differences between academic units as some soft constraints are important for an academic unit and some soft constraints are not. For example, a student would not want to assign more than two exams in the same day. The most common soft constraint in the exam timetabling literature is to spread conflicting exams as much as possible throughout the examination session so that students can have enough revision time between exams (Qu et al., 2009). Other common soft constraints are as follows:

- An exam may need to be assigned before/after the other exam.
- Students should not have exams in following periods.
- Students should not want to take two exams on the same day.
- Same length exams can only be allocated into the same room.
- Exams must (not) be in certain time periods.
- Exams may be split to similar locations.

3. EXAM TIMETABLING APPROACHES

In the timetabling literature, there have been a lot of research using different approaches. Some approaches include combinations of a number of other approaches to improve the solution quality.

This study will concentrate on Tabu Search, Simulated Annealing, Hill Climbing and Great Deluge from Local Search Based Techniques.

Local search is a family of general-purpose techniques for search and optimization problems, which has gain popularity in the AI community (Di Gaspero & Schaerf, 2001). Local search methods are intensively used for optimization problems. In local search methods, one or more initial solution is created using different techniques first. Then these methods move one solution to another one that is inside its neighbourhood considering different local search techniques until some stopping criteria is reached.

These techniques have generated good solutions, but they have used much computer resources.

Local search techniques differ from each other according to the strategy they use both to select the move in each state and to stop the search. In all techniques, the search is driven by a cost function f that estimates the quality of the state. For optimization problems, f generally accounts for the number of violated constraints and for the objective function of the problem.

Local search based techniques and evolutionary algorithms are often classified as meta-heuristics in the literature. Meta-heuristics have been used since the beginnings of operation research to solve difficult combinatorial problems.

3.1 Tabu Search

Tabu Search is originally defined by Glover (1986) to solve optimization problems that have appeared in the literature. After 1986, Tabu Search is intensively used and generated good and successful solutions. These successes have made Tabu Search extremely popular for combinatorial and optimization problems.

Glover (1986) defines the Tabu Search technique briefly below:

“Tabu Search may be viewed as a “meta-heuristic” superimposed on another heuristic. The approach undertakes to transcend local optimality by a strategy of forbidding (or, more broadly, penalizing) certain moves. The purpose of classing a move forbidden, i.e. “tabu”, is chiefly to prevent cycling. In view of the mechanism adopted for this purpose, the approach might alternatively be called “weak inhibition” search, for the moves it holds tabu are generally a small

fraction of those available, and a move loses its tabu status to become once again accessible after a relatively short time. (In this respect the method may be contrasted to branch and bound, which likewise forbids certain moves to prevent cycling, but in a more rigid fashion—a form of “strong inhibition” search.)”

Tabu Search explores the search space by not re-visiting a list of recent moves (kept in a tabu list) (Qu et al., 2009). The search continues moving to other solutions with the aim of escaping from local optima.

The basic principle of TS is to continue the search whenever a local optimum is encountered by allowing non-improving moves; cycling back to previously visited solutions is prevented by the use of memories, called tabu lists, that records the recent history of the search (Gendreau & Potvin, 2005).

3.2 Simulated Annealing

Simulated Annealing is defined by Kirkpatrick and Vecchi (1983) and independently Černý (1985) to solve combinatorial optimization problems. This technique arises from likeness between the physical annealing process of solids and the problem of solving large combinatorial optimization problems (Aarts et al., 2005).

In physics, annealing is known as a thermal process for obtaining low-energy states of a solid in a heat bath. The process consists of the following two steps:

- Increase the temperature of the heat bath to a maximum value at which the solid melts.
- Decrease carefully the temperature of the heat bath until the particles arrange themselves in the ground state of the solid.

Before using the Simulated Annealing method, many parameters like the initial and final temperatures, neighbourhood structure and the cooling factor in the cooling schedule need to be arranged. These parameters affect the performance and success of this technique.

3.3 Hill Climbing

The Hill Climbing algorithm is the most basic local search technique. The Hill Climbing algorithm starts with an initial solution, x , called the current solution with the objective function $o(x)$ (e.g., the number of violated soft constraints) (Merlot et al., 2003). The algorithm processes each exam in an order. The algorithm selects the best neighbour in the all neighbours. This means the selection that minimizes the objective function.

If there is no better neighbour than the current one, the search sets the current one unchanged and it starts the search from another initial (e.g., a randomly selected neighbour) selection.

If a hybrid (combination of some algorithms) algorithm is used to solve the exam timetabling problem, the Hill Climbing algorithm is used to improve the quality of solution (timetable).

The Hill Climbing algorithm doesn't accept worse moves. Thus, it can search a small part of the search space. Because of selecting at least as good as the current neighbour, it never produces a solution that is worse than the original one. It can be said that this algorithm is faster than the others (Burke & Newall, 2003).

4. EXAM TIMETABLING PROBLEM AT MARMARA UNIVERSITY

Marmara University is one of the largest universities of Turkey that has 12 campuses, 16 faculties, 11 institutes and 9 schools. At the university there are currently total 199 associate degrees and undergraduate programs. There are more than 70,000 students and close to 3,000 academic staffs.

In this study, a faculty (Faculty of Engineering) and a department (Information and Records Management Department (IRM) of Faculty of Arts and Sciences) are selected as an example.

Exam timetables at Faculty of Arts and Sciences are prepared by an academic staff at each department manually. Every department has their own limited number of rooms and laboratories.

If a department needs more than their own rooms, the department can use other departments' rooms if rooms are available. IRM is selected as an example because this department has evening education and two curriculums for each education. So this department's exam timetabling is highly constrained. At this department's 2013-2014 fall semester, there are 103 exams, 498 students, 4 rooms which total exam capacities are 144 and 50 time periods, duration of every time period is 60 minutes. At this department's 2013-2014 spring semester, there are 107 exams, 485 students, 4 rooms which total exam capacities are 144 and 50 time periods, duration of every time period is 60 minutes.

Exam timetables at Faculty of Engineering are prepared considering the whole faculty and preparing exam timetables is a challenging task because of limited number of rooms, limited time periods and other constraints. There are 8 departments and close to 1800 students at the faculty. The faculty has not evening education. There are common courses at the faculty and students from each department can take these common courses. When preparing exam timetables for each department, exams of the common courses are assigned first and then the other exams are assigned. At this faculty's 2013-2014 spring semester, there are 388 exams, 1743 students, 42 rooms which total exam capacities are 1169, 28 time periods for midterm and 56 time periods for final. Duration of time periods is usually 120 minutes. Common courses and usage of common rooms make the exam timetabling problem difficult to solve and to generate a feasible solution including the soft constraints.

4.1 Problem Definition

IRM has two curriculums. At these curriculums there are common courses and exams of these common courses are always assigned to the same time period, sometimes to the same room(s). Exams of normal education and evening education courses are sometimes assigned to the same time period if room capacities are enough.

Faculty of Engineering has 8 departments and each department has its own curriculum. Some courses are common courses that are taken by students from different departments.

There are one week for midterm exams and two weeks for final exams at Marmara University. Weekdays are preferred to use first for assigning exams and if there are not enough time periods at weekdays and then weekend days are used. Durations of time periods change from exam to exam and department to department. For example, duration of exams at IRM are always 60 minutes and duration of exams at Faculty of Engineering are usually 120 minutes. So it would not be correct to specify the total numbers of time periods for midterm and final weeks. Also there are two academic semesters and in each semester there are three exam terms; midterm, final and final make-up. But exam timetables for final make-up term are same as the timetables for final term. So it is not needed to prepare exam timetables for final make-up term. Hence, exam timetables are generated four times at a year.

In an overview of ETP of Marmara University, there are various conditions have to be considered. It is clear that ETP can vary greatly both between and within departments and faculties. There are departments that have only normal education and departments that have both normal and evening education. Also some departments have more than one curriculum. Some faculties are preparing the exam timetables of their departments at faculty level considering the limited number of rooms, on the other hand in some faculties, departments prepare their own exam timetables. So these situations increase the number of hard and soft constraints. Hard constraints must be satisfied to generate feasible exam timetables. Soft constraints should be respected as much as possible to generate high quality exam timetables.

4.2 Data Collection

Faculty of Engineering and IRM data are obtained from the Information Management System of Marmara University. The data are listed at the below:

- **Exam:** For each exam, information about course number and department are obtained.
- **Student:** For each student, information about student number, department and enrolled courses are obtained.

- **Room:** For each room, information about normal seating capacity, exam seating capacity and equipment are obtained.
- **Instructor:** For each instructor, information about unique instructor number and department are obtained.

4.3 Data Formats

In this part, detailed information about the data formats are presented. Examples are given for each data formats.

4.3.1 Exam Data Format

| Field Name | Value | Description |
|------------|----------|--------------------------------------|
| ID | 183036 | Unique ID of the exam that is in IMS |
| Name | CSE396.1 | Name of the exam |
| Length | 120 | Length of the exam in minutes |
| Alt | True | Seating type (Full/Exam) |

Alt field specifies the seating type of an exam. If alt field is true, it means that exam capacity of the rooms are used. If alt field is false, it means that full capacity of the rooms are used.

4.3.2 Student Data Format

| Field Name | Value | Description |
|------------|-----------|-----------------------|
| ID | 150110058 | Unique student number |

4.3.3 Room Data Format

| Field Name | Value | Description |
|---------------|----------|--------------------------------------|
| ID | 21 | Unique ID of the room that is in IMS |
| Name | GZMB.345 | Name of the room |
| Full Capacity | 55 | Full capacity of the room |
| Exam Capacity | 28 | Exam capacity of the room |

If the seating type of an exam is set to true, it means that exam capacity of the room will be used in the application.

4.3.4 Period Data Format

| Field Name | Value | Description |
|----------------|----------------|----------------------------------|
| ID | 1 | Unique ID of the period |
| Duration (Min) | 120 | Length of the period in minute |
| Day | Mon 05/26 | Day of the period |
| Time | 9:00a - 11:00a | Start and end time of the period |
| Penalty | 0 ... 4 | The value for priority |

Penalty value ranges from 0 to 4 to determine which periods will be used primarily. If penalty value is 0, it means that the period will be used primarily. If the penalty value is 1, it means that the period will be used after all periods that penalty value is 0. The periods that penalty value is 4 will be used last.

4.3.5 Constraint Data Format

| Field Name | Value | Description |
|------------|-------------|-----------------------------|
| ID | 1 | Unique ID of the constraint |
| Name | same-period | Name of the constraint |

There are four constraint; same period, different period, same room and different room. These constraints can be used to specify the exams.

4.3.6 Parameter Data Format

| Field Name | Value | Description |
|------------|-----------------|------------------------|
| Name | Direct Conflict | Name of the parameter |
| Value | 100 | Value of the parameter |

The parameters are one of the important parts of the application. Parameters that are used in the application can be defined like as in the table. Value field can be integer, long, double and string. If parameters are not defined before the application runs, the predefined parameters will be used in the application.

5. TEST RESULTS

5.1 Test Datasets

One department, IRM, from Faculty of Arts and Sciences and one faculty, Faculty of Engineering, are selected as an example.

5.1.1 Information and Records Management Department

This department is selected as an example because it has evening education and two curriculums for each education. Also all exams of evening education must be after 14:00 pm. Midterm week is one week and final week is two weeks. So this department's ETP is highly constrained. This department's 2013-2014 fall semester features are listed in below:

| | |
|-------------------------------|---------------------------|
| Number of Exams: | 107 |
| Number of Students: | 485 |
| Number of Periods: | 50 (Midterm) / 70 (Final) |
| Number of Rooms: | 4 |
| Total Room Capacities: | 304 |
| Exam Capacities of All Rooms: | 144 |
| Period Length: | 60 minutes |

5.1.1 Faculty of Engineering

This faculty is selected as an example because exam timetabling of this faculty is prepared considering the whole faculty. The faculty has 8 departments. It has not evening education. Midterm week is one week and final week is two weeks. This faculty's 2013-2014 spring semester features are listed in the below:

| | |
|-------------------------------|---------------------------|
| Number of Exams: | 388 |
| Number of Students: | 1743 |
| Number of Periods: | 28 (Midterm) / 56 (Final) |
| Number of Rooms: | 42 |
| Total Room Capacities: | 1907 |
| Exam Capacities of All Rooms: | 1169 |
| Period Length: | 120 minutes |

5.2 Tests

In the literature, there are two options for measuring the quality of solutions; spent time and violated constraints in the solutions. Generally, the quality of timetables is measured by checking to what extent the soft constraints are violated in the solutions generated. In this thesis, this approach is selected. Thus, a weight and penalty value is given to each criteria and total cost of the solution is calculated. Evaluating the total cost of the solution and all criteria values informs about the quality of the solution.

In the tests, selected meta-heuristics are first implemented individually and then together. All these meta-heuristics generate their initial solutions. Test results are briefly explained in the below.

5.2.1 Comparison of Meta-Heuristics on IRM Spring Semester–Midterm

In the table, similar exams are not combined at the same room and the results of this situation is presented in the below. As shown, Hybrid gives better result than the others.

| | Total Exams | Assigned Exams | Perc. (%) | Total Student | More Than n A Day | Direct Conflicts | Back to Back | Total Cost |
|--------|-------------|----------------|-----------|---------------|-------------------|------------------|--------------|------------|
| TS | 107 | 94 | 87.85 | 485 | 82 | 30 | 104 | 39852.05 |
| SA | 107 | 89 | 83.18 | 485 | 90 | 2 | 59 | 12349.84 |
| HC | 107 | 85 | 79.44 | 485 | 89 | 2 | 62 | 12277.60 |
| GD | 107 | 90 | 84.11 | 485 | 180 | 2 | 103 | 21790.16 |
| Hybrid | 107 | 94 | 87.85 | 485 | 80 | 6 | 136 | 15939.27 |

5.2.2 Comparison of Meta-Heuristics on IRM Spring Semester–Final

In the final week, there are more periods than the midterm week. Thus, similar exams are not combined at the same room. There are less room constraints (exams at the same room) than the midterm week.

As shown in table, Hybrid gives both best results of the total cost and the number of direct conflicts.

| | Total Exams | Assigned Exams | Perc. (%) | Total Student | More Than n A Day | Direct Conflicts | Back to Back | Total Cost |
|--------|-------------|----------------|-----------|---------------|-------------------|------------------|--------------|------------|
| TS | 107 | 101 | 94.39 | 485 | 1 | 1 | 23 | 2059.62 |
| SA | 107 | 101 | 94.39 | 485 | 1 | 0 | 16 | 1036.51 |
| HC | 107 | 93 | 86.92 | 485 | 3 | 0 | 18 | 1246.66 |
| GD | 107 | 101 | 94.39 | 485 | 5 | 2 | 50 | 3775.34 |
| Hybrid | 107 | 101 | 94.39 | 485 | 4 | 0 | 15 | 1219.87 |

5.2.3 Comparison of Meta-Heuristics on Faculty of Engineering Spring Semester–Midterm

Exam timetables of the faculty of engineering are prepared as a whole. Thus, there are more exams, students and rooms.

As shown in table, Hybrid gives both best results of the total cost and the number of direct conflicts. In all tests, there are two unassigned exams because of the maximum number of rooms is 4. These two exams are big exams and are needed more than 4 rooms. If maximum number of rooms is set to 10, all exams are assigned.

| | Total Exams | Assigned Exams | Perc. (%) | Total Student | More Than n A Day | Direct Conflicts | Back to Back | Total Cost |
|--------|-------------|----------------|-----------|---------------|-------------------|------------------|--------------|------------|
| TS | 388 | 386 | 99.48 | 1743 | 61 | 41 | 179 | 54023.93 |
| SA | 388 | 386 | 99.48 | 1743 | 36 | 43 | 92 | 52720.78 |
| HC | 388 | 386 | 99.48 | 1743 | 44 | 44 | 162 | 55126.96 |
| GD | 388 | 386 | 99.48 | 1743 | 69 | 44 | 223 | 58401.60 |
| Hybrid | 388 | 386 | 99.48 | 1743 | 42 | 40 | 137 | 50746.99 |

5.2.3 Comparison of Meta-Heuristics on Faculty of Engineering Spring Semester–Final

In the final week, there are more periods than the midterm week too. As shown in table, Hybrid gives the best results again. There are still two unassigned exams because of the maximum number of rooms in which exams assigned.

| | Total Exams | Assigned Exams | Perc. (%) | Total Student | More Than n A Day | Direct Conflicts | Back to Back | Total Cost |
|--------|-------------|----------------|-----------|---------------|-------------------|------------------|--------------|------------|
| TS | 388 | 386 | 99.48 | 1743 | 14 | 91 | 22 | 94352.18 |
| SA | 388 | 386 | 99.48 | 1743 | 5 | 41 | 5 | 43020.83 |
| HC | 388 | 383 | 98.71 | 1743 | 4 | 39 | 13 | 41545.80 |
| GD | 388 | 386 | 99.48 | 1743 | 17 | 41 | 77 | 46078.40 |
| Hybrid | 388 | 385 | 99.23 | 1743 | 4 | 38 | 4 | 39857.95 |

6. CONCLUSIONS

Timetabling problems are combinatorial optimization problems and one of the most effective solving methods of these problems is to use meta-heuristics. In this study, four meta-heuristics from local search based techniques were selected and implemented individually and together. Each of selected meta-heuristics has their own parameter values. For using these meta-heuristics effectively, fine-tuning the parameters of the meta-heuristics is an important step of using these meta-heuristics effectively. All meta-heuristics utilized their best effective parameter values. Besides, neighbourhood selection techniques of the meta-heuristics are another important step.

In the tests, it is needed to pay attention following situations evaluating the results correctly. One of these situations is the success of the meta-heuristics and their hybrid usage when number of time periods and rooms are inadequate. This situation occurred at midterm timetables results of the department of IRM at fall and spring semesters. The meta-heuristics and their hybrid usage gave close results, but hybrid usage was better than others. The other situation is the success of the meta-heuristics and their hybrid usage when number of time periods and rooms are adequate. This situation occurred at both final exam timetables results of the department of IRM and Faculty of Engineering fall and spring semesters. The power and the effectiveness of the hybrid usage of selected meta-heuristics were observed in this situation.

It is suggested that a separate study on fine-tuning the parameters of the meta-heuristics and selecting the appropriate neighbourhood selection techniques can be done.

Another object of this study is to prepare exam timetables faster and error free. Therefore, it is anticipated that this study will contribute to universities and other institutions directly or indirectly. In addition to educational timetabling contributions, this study contributes to the scheduling and timetabling literature. Also this study will be included into the Information Management System of Marmara University and exam timetables will be prepared by this system.

7. REFERENCES

- [1] Aarts, E., Korst, J., & Michiels, W. (2005). Simulated annealing. In E. K. Burke & G. Kendall (Eds.), *Search methodologies* (pp. 187-210): Springer.
- [2] Burke, E. K., De Causmaecker, P., Berghe, G. V., & Van Landeghem, H. (2004). The state of the art of nurse rostering. *Journal of Scheduling*, 7(6), 441-499.
- [3] Burke, E. K., de Werra, D., & Kingston, J. (2004). Applications to timetabling. In J. Gross & J. Yellen (Eds.), *The handbook of graph theory* (pp. 445-474): Chapman Hall/CRC.
- [4] Burke, E. K., & Newall, J. P. (2003). Enhancing timetable solutions with local search methods *Practice and Theory of Automated Timetabling IV* (pp. 195-206): Springer.
- [5] Burke, E. K., & Petrovic, S. (2002). Recent research directions in automated timetabling. *European Journal of Operational Research*, 140(2), 266-280.
- [6] Černý, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of optimization theory and applications*, 45(1), 41-51.
- [7] Di Gaspero, L., & Schaerf, A. (2001). Tabu search techniques for examination timetabling *Practice and Theory of Automated Timetabling III* (pp. 104-117): Springer.
- [8] Easton, K., Nemhauser, G., & Trick, M. (2004). *Sports Scheduling*. In J. Leung (Ed.), *Handbook of scheduling: algorithms, models, and performance analysis*: CRC Press.
- [9] Gendreau, M., & Potvin, J.-Y. (2005). Tabu search *Search methodologies* (pp. 165-186): Springer.
- [10] Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, 13(5), 533-549.
- [11] Kirkpatrick, S., & Vecchi, M. (1983). Optimization by simulated annealing. *science*, 220(4598), 671-680.
- [12] Kwan, R. (2004). *Bus and Train Driver Scheduling*. In J. Leung (Ed.), *Handbook of scheduling: algorithms, models, and performance analysis*: CRC Press.
- [13] Merlot, L. T., Boland, N., Hughes, B. D., & Stuckey, P. J. (2003). A hybrid algorithm for the examination timetabling problem *Practice and theory of automated timetabling IV* (pp. 207-231): Springer.
- [14] Qu, R., Burke, E. K., McCollum, B., Merlot, L. T., & Lee, S. Y. (2009). A survey of search methodologies and automated system development for examination timetabling. *Journal of Scheduling*, 12(1), 55-89.
- [15] Schaerf, A. (1999). A survey of automated timetabling. *Artificial intelligence review*, 13(2), 87-127.
- [16] Wren, A. (1996). Scheduling, timetabling and rostering—a special relationship? *Practice and theory of automated timetabling* (pp. 46-75): Springer.

RSRP: Risk Sensitive Routing Protocol in Wireless Sensor Networks

M Karakaya¹

¹ Department of Computer Engineering, Atilim University,
Incek, Ankara, TURKEY

[¹murat.karakaya@atilim.edu.tr](mailto:murat.karakaya@atilim.edu.tr)

ABSTRACT

In Wireless Sensor Networks (WSN), data communication is mostly implemented by multi-hop transfers of the messages among sensor nodes and the sink. However, due to the nature of wireless communications, WSNs are subject to various malicious attacks. There are various proposals to assess the risk of the network links. In this paper we propose an routing algorithm to exploit the risk assessment of the links to create risk sensitive routes. We have implemented the proposed solution adapting the Ant Colony Optimization method. The simulation results support the success of the proposed method under different network conditions.

Keywords – Security, Wireless Sensor Networks, routing, Ant Colony Optimization

1. INTRODUCTION

Wireless Sensor Networks (WSNs) get high popularity due to several important properties such as easy deployment, self maintenance, low energy spending, high resistance to node failures, etc. These properties mostly depend on the underlying communications method. One of the mostly implemented methods is to pass the messages between source and destination node via intermediate nodes. That is, source and destination nodes do not have a direct link; rather, they rely on the intermediate nodes forwarding the messages. In general, this way of communication is called multi-hop communication. The advantages of the multi-hop communication are dynamic routing, robust to node failures due to energy or malicious attacks, ease of deployment and management.

However, WSNs has their own drawbacks as well. One of the important drawbacks is the sensitivity of the communication links among nodes. An attacker can attack to these wireless connections among the nodes such that some nodes cannot communicate with the sink [3]. The most routing algorithms used in WSNs are adapted from the Ad-Hoc Wireless networks [1][2]. The high priority design challenges of the proposed routing protocols are usually Minimal computational and memory requirements, self-organization, energy efficiency, scalability, etc. [2]. Thus, most of these routing protocols do not directly consider the security treats directly [3]. Actually, there are many different kinds of security treats based on wireless communications: jamming, flooding, tampering or destruction, continuous channel access, replaying existing messages, Sybil and wormhole attacks, etc. [3].

One solution to provide security in routing can be Trust Management System (TMS). Trust, or the trust on the behavior of the elements of the network, is a key aspect for WSN. A trust management system can distinguish a node that does not function as expected due to either faults or malicious acts [4][5]. There have been many different proposals for TMS. TMS can be used to improve the routing security of WSNs if a sensor node knows which other nodes in the network can be trusted for forwarding a packet [4][5]. Thus, nodes can use other nodes reputation to decide which one of these to trust. Thus, TMS evaluate the node operations in WSN and assign each of them with some reputation value and broadcast these reputations.

In this work, we focus on using reputations of nodes to create a path with minimum risk sensitivity. The proposed algorithm aims to generate a routing path such that the total reputation of

the intermediate nodes between the source and the sink node has the lowest security risk. The solution is based on the Ant Colony Optimization [6].

2. RISK SENSITIVE ROUTING

Assume that a TMS has been implemented to observe nodes in a WSN and produces some reputations per nodes as shown in Fig. 1. For ease of interpretation, we suppose that TMS determines the risk of a given link between two nodes rather than a node's risk. For instance, sending a message from Node 1 to 2 has a low level of security risk (10%), where transmitting a package from Node 7 to sink has a high level of security risk (60%). Using the given topology and the risk assessments in Fig. 1, we can create many different routes with different number of hops and cumulative risk values. In Fig. 2, we present 3 possible routes. In Route 1, the path from Source to Sinks via Node 7 has only 2 hops and the cumulative risk would be 68%.

However, if we follow Route 2, hop number would be 4 but the total risk value becomes 34%. As a last example, Route 3 produces 3 hops with an aggregated risk of %35. The bottom line of these examples is the importance of selecting immediate nodes on the number of hops and accumulated risk values. To save intermediate node's energy, most routing algorithms attach high importance to minimizing the hop number between source and target nodes. However, this requirement can conflict with the requirement of low level risk involved in the routing as seen below example.

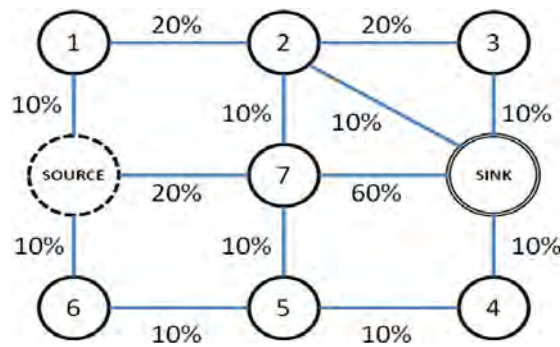


Figure 1. A sample topology with link risk assessments.

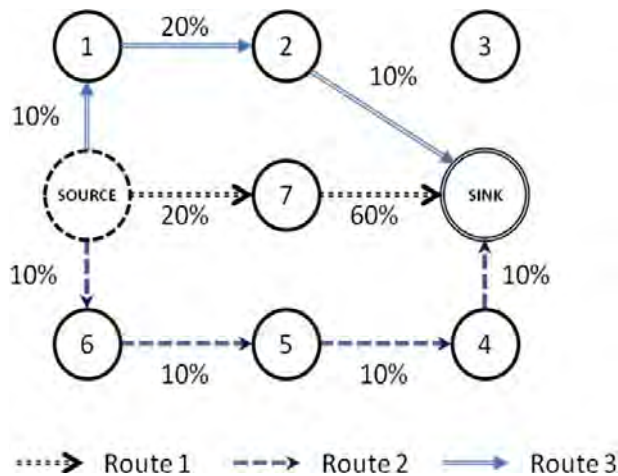


Figure 2. Some of the possible routing paths.

In this work, we provide a Risk Sensitive Routing Protocol (RSRP) such that the trade-off between hop number and risk value of the routing path can be done depending on a system parameter, Risk Sensitivity (RS), defined by the WSN operator. RS shows how much WSN is sensitive to possible lost of packages due to low level of reputations of the nodes in the network. RS can be any value between 0 and 1. When RS is assigned 0 it means that we do not care the

risks at all and we definitely want to minimize the number of hops in the routes. On the other hand, if RS is assigned 1 this states that we do care about the risks in the routes and want to minimize the level of the risks at the cost of increasing number of hops. Any value between 0 and 1 is an intermediate sensitivity to the risks.

We can now define the target function (TF) of the RSRP as Eq.1, where CR stands for Cumulative Risk of the given route, HN is the Hop Number. RSRP aims to minimize the TF value for the given RS value by selecting an appropriate route among possible routes in WSN.

$$TF(R) = (RS * CR) * ((1 - RS) * HN) \quad (1)$$

Using the example in Fig.1 and Fig. 2, if RS is 0, the RSRP would select Route 1, the shortest path between Source and Sink via Node 7. However, if RS is set 1, then the RSRP would select Route 2 locating the path with the lowest cumulative risk. If we set RS 0.5, the RSRP seeks a balance between security and efficiency by selecting Route 3.

3. Ant Colony optimization

Ant Colony Optimization (ACO) is an evolutionary metaheuristic which mimics the ants' foraging behavior to find the optimum solutions of a given problem [6]. ACO has been adapted to many different engineering problems successfully [7]. We have developed several solutions for the routing messages in WSN [8] and scheduling target for Unmanned Aerial Vehicles (UAV) [9, 10]. In this work, we apply ACO to optimize the TF of possible routes. Details of the implementation are given below.

3.1 Finding Next Node

In ACO, each artificial ant tries to create a routing path by visiting nodes bearing in mind the given Risk Sensitivity (RS) and the Target Function (TS) values. Beginning from Source Node, each ant calculates the probability of movement from the current node to the all unvisited nodes as in the following formula:

$$P_{ij} = \frac{(\tau_{ij})(\eta_{ij})^\beta}{\sum_{i,j \in M_k} (\tau_{ij})(\eta_{ij})^\beta}, \quad i, j \notin M_k \quad (2)$$

In the formula, i is the current location, j is the possible next location, τ_{ij} is the pheromone value between two locations, η_{ij} is the heuristic value between two locations, β is the coefficient for the heuristic parameter, and M_k is the memory for storing list of the nodes which are already visited by the ant k . Thus, P_{ij} is the normalized probability of ant k to move from node i to node j . After finding the movement probability for all the nodes, a random number between 0 and 1 is drawn to select the next node according to total probabilities of all the possible nodes. If the selected next node is the Sink, then search progress terminates with a routing path and ant returns to the Source node to restart another route construction. Otherwise, ant continues to move on as described above.

Each ant builds its own route planning simultaneously by exploiting the experiences of other ants by sensing the pheromone values in the formula 2.

3.2 Updating Pheromone Values

After completing a tour, each ant calculates the tour cost as given in Formula 1. Before applying any pheromone update on the targets, first evaporation should take place. Thus, all the pheromone values between all the targets are decreased using the evaporation parameter value (p) as in the following formula:

$$\tau_{ij} = (1 - p)\tau_{ij}, \quad i, j \in H \quad (7)$$

Then, all nodes on the route constructed by the ant (R_k) receive an update depending on the TF value of the tour:

$$\tau_{ij} = \tau_{ij} + \left(\frac{1}{TF}\right), i, j \in R_k \quad (8)$$

Formula 8 dictates that the solutions with less TF, that is causing less number of hops and less amount of security risk, drop more pheromone on the paths to provide positive feedback for the other ants.

3.3 Finding Heuristic Value

The heuristic value (η_{ij}) between two nodes is defined according to the given RS value. Remember that RS value can be between 0 and 1. When RS is assigned 0 it means that we do not care the risks at all and we definitely want to minimize the number of hops in the routes. Thus, the heuristic value (η_{ij}) only depends on the hop number between two nodes. If there is a link between them link existence (l_{ij}) will be 1 and η_{ij} will be set 1 as well. However, if there is no link connecting these two nodes, link existence (l_{ij}) and η_{ij} will be set 0.

On the other hand, if RS is 1, this states that we do care about the risks in the routes and want to minimize the level of the risks at the cost of increasing number of hops. As a result, the heuristic value (η_{ij}) only depends on the risk probability between two connected nodes, which can be calculated as:

$$\eta_{ij} = \frac{1}{r_{ij}} \quad (9)$$

where r_{ij} is the given risk probability of the link between the nodes. If these two nodes are disconnected ($l_{ij}=0$), η_{ij} will be zero as well.

As a last option, RS can take any value between 0 and 1. In this case, the heuristic value (η_{ij}) is found based on the following formula provided that the given two nodes are connected ($l_{ij}=1$):

$$\eta_{ij} = \frac{1}{r_{ij} * RS} \quad (10)$$

3.4 The Proposed RSRP Algorithm

Using the main ACO functions explained above the pseudo code of the RSRP is given in Table 1. We input the target list (H), the link existence between the nodes (l_{ij}), the risk probabilities between the targets (r_{ij}), and the Risk Sensitivity (RS). Then, using these parameters, the heuristic values are calculated. After creating a number of ants (m), each ant builds its solution and updates the pheromone values according to the cost of the solution based on the Formula 1. When, a pre-defined number of iterations has been executed algorithm terminates by outputting the best solution found so far.

In some cases, ant could be in a dead end such that from the current node it cannot move forward. Thus, it should move back to the previous node to see that if it can move to another node other than the dead end node.

4. SIMULATION TESTS AND OBTAINED RESULTS

We have implemented the proposed RSRP using Java. The simulation and ACO parameters with the default values are given in Table 2.

4.1 Network Topology and Risk Distribution

We use an NxN size grid topology with different risk distributions. To assign risk values to the links between nodes we have propose to use a Risk Distribution (RS) set up. We assume that most of the nodes in a WSN is dependable but open to failure. Therefore, their associated risk would be very low. However, the attackers in the network present much more risks in routing the packets. Therefore we assign higher level of risks to the attacker.

As a result of above observation, we separate the links into two groups: friendly and malicious. Furthermore, the risk level posed by the malicious nodes is classified in three levels. In all scenarios explained below we assume that 90% of the links in the grid is friendly and has a risk probability at most 0.1%. However, the rest of the links (10%) has different risk levels according to the risk distribution. In *Low Risk Distribution* (LRD), the malicious links has a risk probability up to 10%, in *Moderate Risk Distribution* (MRD), they have a risk probability up to 50%. In the last distribution, *High Risk Distribution* (HRD), risk probabilities of malicious links are 80% at most.

We select the first node as the sink and the node in the farthest corner of the grid as the source node. As an illustration, in Fig.1, Node 1 is the sink; Node 4 is the source node.

Table 1. *RSRP pseudo code.*

```
RSRP ( $H, l_{ij}, r_{ij}, RS$ )
{
  init_Pheromone_Values();
  init_Heuristic( $M_{ij}$ );
  create_Ants( $m, Source$ );

  while (!end_condition_satisfied)
  {
    for each ant
    {
      while (Sink has not been reached)
      {
        next = find_Next_Target();
        if (Sink_Reachable(next))
        {
          move_forward(next);
        }
        else {
          move_backward();
        }
      } //end_while
      evoporate_Pheromone();
      update_Pheromone();
      update_Best_Solution();
    } //end_for_each_ant
  }
  return (Best_Solution);
}
```

Table 2. *Simulation parameters.*

| Parameter | Definition | Default Value | Notes |
|-----------|-------------------------|---------------|------------------------------|
| N | Grid Size | 10 | Total number of nodes is 100 |
| RS | Risk Sensitivity | - | 0, 0.5, or 1 |
| RD | Risk Distribution | | Low, Moderate, or High |
| AN | Ant Number | 400 | |
| IN | <i>Iteration Number</i> | <i>4000</i> | |
| P | Evaporation Rate | 1% | |
| β | Heuristic Coefficient | 2 | |

For the results, we have run each tests 10 times and reported the mean values in the tables below. Moreover, we have tested the RSRP for the three different source nodes and provided the average results as follows.

4.2 Results for Low Risk Distribution

When the links in the grid topology is initiated according to the Low Risk Distribution (LRD) as described above, source node has the opportunity to use many different routing paths. In Table 3, we observe the effect of the RSRP according to the given risk sensitivity (RS). If we do not care about the risk (RS=0) and would like to have only a shortest routing path, the RSRP is able to create the shortest path with 58 hops (HN). However, the cumulative risk (CR) turns out to be around 17%.

Table 3. *Results for LRD.*

| RS | HN | CR (%) |
|-----|----|--------|
| 0 | 58 | 16.85 |
| 0.5 | 66 | 2.53 |
| 1 | 70 | 2.38 |

On the other hand, if we only care about the CR, we can set RS 1. As the results in Table 3 show the RSRP builds a routing path with CR = 2.38 and HN = 70. Thus, we find a path with a very low risk but longer than the shortest one.

As a balance between these two choices, we can assign 0.5 to RS and gain the results as HN= 66 and CR=2.53. This results support the balance between shortest and lowest risk paths clearly.

As a result of all the observations in Table 3 indicate that the RSRP functions successfully as designed for.

4.3 Results for Moderate Risk Distribution

As seen in Table 4, increasing the risk stemmed from the attackers cause a drastic change in the level of cumulative risk of the shortest path. However, the RSRP is still able to find routing paths walking around the attackers successfully when RS is set either 1 or 0.5. Actually, the HN and CR are almost the same as the ones of the LRD.

Table 4. Results for MRD.

| RS | HN | CR (%) |
|-----|----|--------|
| 0 | 58 | 54.86 |
| 0.5 | 64 | 2.46 |
| 1 | 70 | 2.25 |

4.4 Results for High Risk Distribution

As a last option we again increment the attackers' risk level up to 80%. The RSRP can successfully build routes with low CR values when RS is different than 0. However, when RS is 0, the generated shortest path has very high level of risk as much as 75%.

Table 5. Results for HRD.

| RS | HN | CR (%) |
|-----|----|--------|
| 0 | 58 | 75,40 |
| 0.5 | 72 | 2,44 |
| 1 | 85 | 2,33 |

4.5 Discussion

In the simulation tests, we have observed that the RSRP is sensitive to risks in building routing paths. Even in high level of risks attached to some links, the RSRP finds paths excluding these risky links at a cost of increasing the total number of hops.

Furthermore, by providing an important parameter, risk sensitivity (RS), the RSRP enables the user to decide the balance between cumulative risk on the routing path and the length of the route. In the experiments we have observed that instead of minimizing risk by setting RS 1, we can realize shorter paths with very low cumulative risk by assigning 0.5 to RS.

5. CONCLUSIONS

In this work, we define a practical problem faced in routing in multi-hop communication systems such as WSNs. The aim is to take into consideration of the link security risks while planning hops. We propose to adapt the ACO meta-heuristic to solve this problem. Thus, we developed the RSRP in which we can tune the risk sensitivity of the routing protocol and accordingly construct routing paths. After implementing the proposed solution we have conducted various tests to observe its performance under different risk level. The obtained results clearly indicate the success of the RSRP.

6. REFERENCES

- [1] Prathap, U., Shenoy, D.P., Venugopal, K.R., Patnaik, L.M., "Wireless Sensor Networks Applications and Routing Protocols: Survey and Research Challenges," Cloud and Services Computing (ISCOS), 2012 International Symposium on , vol., no., pp.49,56, 17-18 Dec. 2012.
- [2] Muhammad Saleem, Gianni A. Di Caro, "Muddassar Farooq, Swarm intelligence based routing protocol for wireless sensor networks: Survey and future directions", Information Sciences, Volume 181, Issue 20, 15 October 2011, Pages 4597-4624, ISSN 0020-0255

- [3] Kavitha, T., and D. Sridharan. "Security vulnerabilities in wireless sensor networks: A survey." *Journal of information Assurance and Security* 5.1 (2010): 31-44.
- [4] Lopez, Javier, et al. "Trust management systems for wireless sensor networks: Best practices." *Computer Communications* 33.9 (2010): 1086-1093.
- [5] Mármol, Félix Gómez, and Gregorio Martínez Pérez. "Providing trust in wireless sensor networks using a bio-inspired technique." *Telecommunication systems* 46.2 (2011): 163-180.
- [6] Alberto Colomi, Marco Dorigo, Vittorio Maniezzo, et al. "Distributed optimization by ant colonies", In *Proceedings of the first European conference on artificial life*, volume 142, pages 134–142. Paris, France, 1991.
- [7] Chandra Mohan, B., and R. Baskaran. "A survey: Ant Colony Optimization based recent research and implementation on several engineering domain." *Expert Systems with Applications* 39.4 (2012): 4618-4627.
- [8] Murat Karakaya, "Time-Sensitive Ant Colony Optimization to Schedule A Mobile Sink for Data Collection in Wireless Sensor Networks", *The Ad Hoc & Sensor Wireless Networks*, In press, 2014.
- [9] Murat Karakaya, "UAV Route Planning For Maximum Target Coverage", *Computer Science & Engineering: An International Journal (CSEIJ)*, Vol. 4, No. 1, February 2014.
- [10] Murat Karakaya, "Route Planning For Using Minimum Number of UAVs in Observing Fixed Targets", (in Turkish), 15. Otomatik Kontrol Ulusal Toplantısı ve Sergisi (TOK2013), Eylül 2013.

Using Service Oriented Architecture for Plate Recognition by Mobile Devices

M Karakaya¹, G Şengül²

^{1,2}Department of Computer Engineering, Atılım University,
İncek, Ankara, TURKEY

¹*murat.karakaya@atilim.edu.tr*, ²*gokhan.sengul@atilim.edu.tr*

ABSTRACT

Automatic license plate recognition is the process of determining vehicle number plates from images. The process is essentially based on the image processing techniques. In this study we propose a Service Oriented solution for automatic license plate recognition on mobile devices. The main aim of the proposed system is to be used to detect incorrectly parked cars and their owners in special areas such as university campuses. In the proposed system the images are captured by the mobile devices, and those images are transferred to a server by internet or intranet environments. All the image processing applications are done on the server site and finally the server returns the owner's name and his information about the cars to the mobile device. For the plate recognition we used the Speeded Up Robust Features (SURF) approach. The proposed system gives promising results.

Keywords image processing, plate recognition, SURF, SIFT

1. INTRODUCTION

Automatic license plate recognition (ALPR) is the process of determining vehicle number plates from images captured from one or more cameras. The process is essentially based on the image processing techniques. The first step in determining the vehicle plate numbers is the determination of the plate region in the vehicle image, and the second step is to determine the characters on the plates by using the Optical Character Recognition (OCR) algorithms.

License plate recognition systems are commonly used in areas where authorized entry and exit are permitted. Besides they are used to find the stolen vehicles and/or lost vehicles and they are also used for security purposes. License plate recognition systems are probably mostly used in traffic applications nowadays such as traffic law enforcement and road traffic monitoring systems [1-3]. The other application areas of the license plate recognition systems are as follows: parking areas, entry and exit points of highways, bridges, boarders, highway control points, entry points of the sites, entry points of the universities, hospitals, hotels, factories, airports, shopping malls, etc [1-4]. License plate recognition systems are also known as automatic vehicle identification, car plate recognition and automatic number plate recognition [5].

There are many algorithms suggested for automatic number plate recognition systems in the literature. Most of the algorithms are based on the following steps:

Step1: Detection of plate region in the image

Step 2: Resizing the plate for the other steps

Step 3: Normalization of illumination and contrast values of the plate image

Step 4: License plate segmentation

Step 5: Character recognition based on the optical character recognition (OCR) techniques

Step 6: Preprocessing and control checks.

The preliminary step is the image acquisition step in which the images of the license plate are acquired. This step is common for all the license plate acquisition systems and due to this reason it is not included above. The literature survey dealing with the license plate recognition systems can be found in the literature [5] and it is not repeated in here.

The performance of the automatic license plate recognition systems may vary depending on the application areas and environmental conditions. For example, a license plate recognition system set up in a car park entrance may have a higher performance since it is worked in a controlled environment, in which the lighting conditions are fixed and vehicles may have lower speeds during the imaging phase. On the other hand the systems used in highways may have poor performance since the environmental conditions may vary during the imaging (i.e. lighting and weather conditions, etc.) and the vehicles may have higher speeds which may cause blurring effects in the images. In general the state of the art automatic number plate recognition systems have a performance of around 95% for plate detection, and 93% for plate recognition phases [5]. In order to have a high performance it is necessary to use a specific algorithm, specific image capturing devices and network infrastructure for the application.

There are some problems encountered in number plate recognition systems. The problems can be itemized as follows: camera localization, distance of camera to the imaged vehicles, low resolution imaging capability of some cameras, lighting conditions of the imaging environment, contrast and illumination problems and shadow effects, unable to capture the image of the number plates, i.e. due to dust etc., and different font style of the plates [6-7].

In this study we proposed a mobile application that can capture the images of number plates of vehicles, send that image to a server and then return the necessary information about the vehicle such as the owner name of the vehicle and his contact information.

2. PROBLEM DEFINITION

Consider a university campus in which hundreds or thousands of people enter and park their cars. University administration would like to locate the owner of any car in some situations. In most cases, faculty members, students, and administrative personnel have registered their car plate and identity to some car information system located at campus gates. Visitors are also requested to register to the some information system as they arrive in as well. However, when people park and leave their cars, security officers can only locate the owner of a car by calling the gate and exchanging the plate number and identification of the owner, which can be time consuming and this process may require human power and effort. This process may increase the work load of the security officers. So it may be necessary to propose an automated system that will work easily and provide necessary information of the cars parked on the campus.

3. THE PROPOSED SOLUTION AND IMPLEMENTATION DETAILS

Mobile devices such as smart phones, tablet pc, etc., have gained high popularity in real life applications to ease the performing daily jobs. However, due to their limited resources, e.g. battery, computation power, memory, there are novel challenges in applying mobile solutions. In this paper, we propose a Service Oriented solution for image processing on mobile devices and partial implementation of it.

In this work, we offer using a mobile application to remove any human intervention for automated license plate recognition. In the proposed mobile solution, security officers carry a smart phone connected to campus intranet. They use the mobile application to take a picture of the plate of the car and send it to the central server for identification and recognition of the plate number. The central server applies image processing techniques to extract and recognize the plate number. After determining the number, the central server communicates with the registration information system and pulls up the identification details of the owner and sends them to the mobile application on the smart phone.

In the proposed service-oriented architecture, the heavy work is done by the central server. Thus, we can save mobile device's resources. In the work, we implemented the image processing as a service at the central server. Thus, mobile application uses this service for the identification details by sending a plate picture. The central server has two main modules: Image Processing and Identification Retrieval. In the Image Processing module, we identify and extract the license plate number as text from the given picture. In the Identification Retrieval module, we query the car registration database with this plate number to find the owner's details. Then, the central server sends this piece of information to the requesting mobile device. In this paper we focus on the Image Processing module and test its performance. The block diagram of the proposed solution is given in Figure 1. The details are given below.

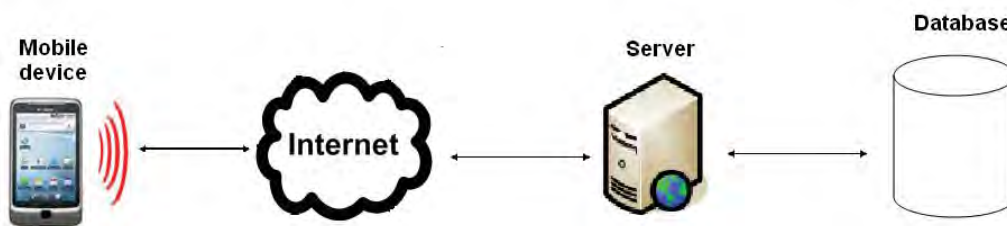


Figure 1- The block diagram of the proposed system. In this system the mobile device is used as the image capturing unit and it also shows the information sent from the server. Image processing and Identification Retrieval modules are run on the server. In the database the information about the registered cars (i.e owner of the car and his/her contact information) are kept.

There are several different approaches to recognize the plate number [5]. However, in these works they mostly assume that car is moving and a video recorder record the moving car. Thus, they try to extract the frames including plate number. However, in our case, car is stationary and we have a picture of the plate. Therefore, our approach is much more direct. First of all, we created an image database including plates printed by various fonts. We collected these images from the cars park in the Atilim University campus. Each image is cropped such that each single image in the database holds only a single number or a character. Then using Matlab [8], we extracted their features according to Scale-invariant feature transform (SIFT) [9] and Speeded Up Robust Features (SURF) [10] algorithms. Thus we have a feature database for numbers and characters printed on plates in various fonts. After then, when a plate image is input, the proposed algorithm first identifies the plate in the picture. The plate detection is performed according to the proposed methods in the literature [11]. In Figure 2 a sample car image containing the license plate is shown and in Figure 3 the plate region identified is depicted.



Figure 2 – A sample car image containing the license plate. The image was acquired in daylight and when the car was stationary.



Figure 3 – The plate region of the car in Figure 2. The characters are shown as black lines.

This section of the image is cropped to investigate further. Calculating the distances between characters and numbers in the cropped image we split the image such that each new image holds a number or a character only. In the last step of the algorithm we calculate the match points between the given number or character and the training set according to SIFT and SURF. Using a voting scheme, we categorize the image and integrate the resulting plate. The character recognition is performed by matching the SURF feature points. We used feature point matching approach previously for logo detection [12]. The matching scheme is based on the minimum distance measurements. In Figure 4 a SURF feature points of training image, test image and matched points are shown.

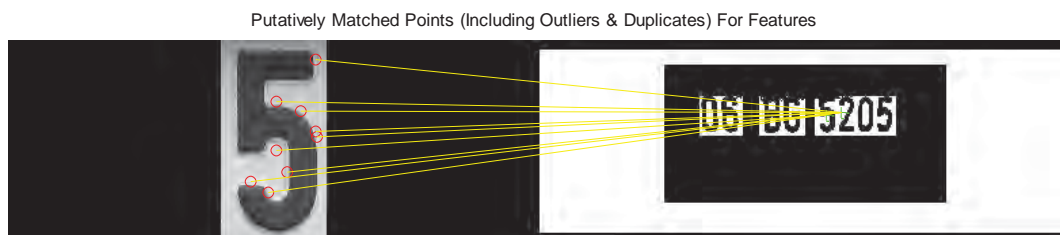


Figure 4- Surf feature points of number 5 and the matched features of the plate. The red circles show the feature points of 5, and green signs show the feature points of the plate region matched by the matching procedure.

4 RESULTS

We have tested the proposed method for 5 test plates and we obtained the correct plate numbers. The results show that the proposed method works correctly. But it is important to note that the tests were performed in a controlled environment and we need to perform more tests to evaluate the performance of the system. The promising results support the success and feasibility of the proposed architecture.

4 DISCUSSIONS AND CONCLUSION

In the experiments of the proposed method, we observed that we can extract the correct plate at a very high probability. We think that it is due to using two important feature extraction algorithms together to decide a match.

We performed the tests for 5 plates. In order to evaluate the performance of the system it is necessary to repeat the tests with a high number of plates. Another important point about the tests is that we performed the tests on a controlled environment, in which the lighting conditions were satisfactory and the cars were stationary. In case of pure light, it can be necessary to work on the algorithm and the performance may vary depending on the lighting conditions. On the other hand another factor that may affect the performance is the weather conditions. It is necessary to repeat that the tests on different conditions.

The proposed system works on the condition in which the imaged cars are not moving. This condition prevents the motion blur effect on the imaging system, which may decrease the performance of the system. The motion blur is reduced by preprocessing the acquired image by applying sharpening filters. The lack of motion in proposed solution is an advantage of the solution.

As a future work, we would like to use a database with a high number of car records so that an integrated solution would be ready to be used in Atılım University.

Acknowledgements: This work is supported by Atılım University by the project number ATU-LAP-C-1314-07.

5 REFERENCES

- [1] G. Liu, Z. Ma, Z. Du, and C. Wen, "The calculation method of road travel time based on license plate recognition technology," in Proc. Adv. Inform. Tech. Educ. Commun. Comput. Inform. Sci., vol. 201. 2011, pp. 385–389.
- [2] S. Kranthi, K. Pranathi, and A. Srisaila, "Automatic number plate recognition," Int. J. Adv. Tech., vol. 2, no. 3, pp. 408–422, 2011.
- [3] C.-N. E. Anagnostopoulos, I. E. Anagnostopoulos, I. D. Psoroulas, V. Loumos, and E. Kayafas, "License plate recognition from still images and video sequences: A survey," IEEE Trans. Intell. Transp. Syst., vol. 9, no. 3, pp. 377–391, Sep. 2008.
- [4] C.N.E. Anagnostopoulos, I.E. Anagnostopoulos, V. Loumos, E. Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications", IEEE Trans. Intell. Transp. Syst. 7, 377–391, 2006.
- [5] S. Du, M. Ibrahim, M. Shehata, "Automatic License Plate Recognition (ALPR): A State-of-the-Art Review, IEEE Transactions on circuits and systems for video technology, vol. 23, No:2, 2013.
- [6] C. Nelson, K. Babu and K. Nallaperumal, "An efficient geometric feature based license plate localization and recognition," Int. J. Imaging Sci. Eng., vol. 2, no. 2, pp. 189–194, 2008.
- [7] D. Zheng, Y. Zhao, and J. Wang, "An efficient method of license plate location," Pattern Recognit. Lett., vol. 26, no. 15, pp. 2431–2438, 2005.
- [8] MATLAB, <http://www.mathworks.com/>, 2014
- [9] Lowe, David G. (1999), "Object recognition from local scale-invariant features", Proceedings of the International Conference on Computer Vision 2. pp. 1150–1157
- [10] H. Bay, A. Ess, Tinne Tuytelaars, Luc Van Gool "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346–359, 2008
- [11] N. B. Gohil, "Car License Plate Detection", Master thesis, California State University, Sacramento, 2010.

[12] M. Karakaya and G. Şengül, "Determining Firm Logos Using Image Processing Techniques on a Mobile Device to Aid Blind People", ISEAI: Second International Symposium on Engineering, Artificial Intelligence and Applications, 5-7 November, 2014.

Determining Firm Logos Using Image Processing Techniques on a Mobile Device to Aid Blind People

M Karakaya¹, G Şengül²

^{1,2}Department of Computer Engineering, Atılım University,
Incek, Ankara, TURKEY

¹murat.karakaya@atilim.edu.tr, ²gokhan.sengul@atilim.edu.tr

ABSTRACT

Many firms associate their brand name with a logo. Thus, customers can face logos and easily recognize the brands without reading their names. Using this strong relationship, we introduce a logo recognition system useful to the blind people. In the proposed system, blind or visually impaired people can use their smart phones as a video recorder in a shopping mall food court to identify the shops around them. The mobile application running on the device captures the frames and analyses them to identify any logos. The identified logo is read to the user. In this paper, we propose a practical yet successful image processing algorithm to enable such a mobile application work on. The initial experiments on sample videos indicate promising results.

Keywords – image processing, blind, logo recognition, logo detection

1. INTRODUCTION

According to the Merriam-Webster Dictionary, logo is “a symbol that is used to identify a company and that appears on its products” [1]. Logo recognition is one of the well known image processing problems faced in many different areas of applications [2,3,4,5,6,7,8,9]. Some researchers use the vehicle logos to describe the vehicle brand [2], some researchers exploit the logos on the printed documents to classify them [3,4,5], and even some researchers benefit logo to display times in the determination of the key moments of a game in a TV broadcast [6].

As the variations in the applications, a number of different image processing techniques have been proposed to solve this problem successfully. In some studies, by using color histogram information of logos which can be considered as colored objects, a method based on edge detection is proposed [5]. Similarly in another study, a method to identify limiting the rectangular shapes in the logo area is proposed [7]. In another study, exploiting low-level gradient channels to identify quickly the logos has been proposed [8]. In order to identify logos from different angles solutions based on SIFT have also been proposed [9]. Similarly for recognition logos in rotated images a further method is based on identifying the logo components are provided in [10].

In this study; we aim to detect and identify the fast food firm logos located on food court of shopping malls to aid blind or impaired-vision people. The problem definition and the details of the solution are presented in the next sections. The data collection and the experiment test are presented in Section 4. Lastly we conclude our remarks in the end of the paper.

2. PROBLEM DEFINITION

Assume that you are blind and would like to have lunch in a nearby shopping center. You arrive at food court and would like to know which fast food companies are around you. Now, you have to ask help from someone and memorize all these shop locations for the next visit. Whenever a new fast food shop opens or whenever you visit a new shopping center you need to ask and memorize again.

Nowadays, mobile devices such as smart phones or tablet PCs are used widely in solving daily life problems. Using the example problem as a motivation, we propose a new application to detect and recognize the logos used in a food court of a shopping center. Thus, the logo recognition problem can be defined as:

“detecting and recognizing the firm logos by using a mobile device and read out their names to a blind or vision impaired person”.

3. THE PROPOSED SOLUTION AND IMPLEMENTATION DETAILS

The method proposed to solve the logo recognition problem has three main phases. In the pre-processing phase, the clip that a mobile device records is examined and frames are extracted as images. In the second phase, these images are processed to extract their features and these features are compared with the features of the logos stored in the logo database. A voting system evaluates the matching of a given image to the known logos. In the last phase of the algorithm, name of the determined logo is read to the user.

The proposed solution is implemented using MATLAB [11]. The details of the solution are provided below.

3.1 Creating Training Logo Database

To have a training set, we have recorded videos of logos at food courts of various shopping malls located at Ankara, Turkey. We have recorded videos from 9 different angles and distance to the logos as shown in Fig. 1. Then, we extract a number of logo images from these videos and label them according to the firm name. After converting colorful frames to grey scale, we have employed *Speeded Up Robust Features* (SURF) [12] technique to detect feature points and extract SURF descriptors in each image. We build an index based on *k-dimensional tree* (KD tree) [13] to store feature vectors of feature points and their descriptors. Thus, we have a logo database holding feature vectors and names of logos. Table 1 provides the MATLAB code details for the explained procedure above.

Table 1. MATLAB code to creating Training Logo Database .

```

for l = 1:numel(logoCollection)
    % Detect SURF feature points
    logoCollection(l).points = detectSURFFeatures(logoCollection(l).image,
        'MetricThreshold',600);
    % Extract SURF descriptors
    [logoCollection(l).featureVectors,logoCollection(l).validPoints] =
        extractFeatures(logoCollection(l).image,logoCollection(l).points);

    % Save the number of features in each image for indexing
    logoCollection(l).featureCount = size(logoCollection(l).featureVectors,1);
end
% Combine all features of logos into dataset
featureDataset = double(vertcat(logoCollection.featureVectors));
% Instantiate KD Tree
imageFeatureKDTree = KDTreeSearcher(featureDataset);

```

3.2 Comparing Captured Frames with the ones Stored in Logo Database

A mobile application records the surroundings and extracts the frames to compare them against the training data set. To do so, we first create a feature vector from the given frame as explained above using the SURF technique. Following that, we compare this vector with the stored vectors in the KD tree. The comparison counts the number of matched feature points between the test and

training images and normalized it according to the total number of feature points in the given vectors. Since we have more than one test image for a single logo we calculate the total matched points for each logo. Thus, we use each test image as a voter. Lastly, if the voting is higher than a threshold value determined by the system, we select the logo label with the highest matching point as the candidate winner. We consider a candidate winner as a final winner if it wins the voting in three consecutive frames.

Table 2. *MATLAB code to comparing captured frames with ones in the training database.*

```
% Detection loop for each frame
for currentFrame = 1:numel(frameCollection)
    query.image=frameCollection(currentFrame).image;
    % Detect SURF features for the queried image
    query.points = detectSURFFeatures(query.image,'MetricThreshold',600);
    % Extract SURF descriptors
    [query.featureVectors,query.points] = extractFeatures(query.image,query.points);
    % Match each query feature to two (K=2) closest features in the dataset.
    [matches, distance] = knnsearch(imageFeatureKDTree,query.featureVectors,'K',2);
    indexIntervals = [0, cumsum([logoCollection.featureCount])] + 1;
    % Ratio Test
    goodRatioMatches = distance(:,1) < distance(:,2) * .8;
    % Distance threshold
    goodDistanceMatches = distance(:,1) < .25;
    goodMatches = matches(goodDistanceMatches & goodRatioMatches,1);
    % Count number of features that matched from each training image using stored
    % indices for dataset matrix
    counts=histc(goodMatches, indexIntervals);
    % Collect the votes
    [~,bestMatchedLogo]=max(counts);
    % Match the features using their descriptors.
    logoPairs = matchFeatures(logoCollection(bestMatchedLogo).featureVectors,
    query.featureVectors);
    % Display matched features.
    matchedlogoPoints = logoCollection(bestMatchedLogo).validPoints(logoPairs(:, 1), :);
    matchedScenePoints = query.points(logoPairs(:, 2), :);
    matchPointNumber = size(matchedlogoPoints);
    if (matchPointNumber(1) < minimumMatchPoint)
        disp('Not detected: Insufficient matched features (Including Outliers & Duplicates)');
        notDetected=notDetected+1;
        continue;
    else
        % Logo detected
        detectedLogo = logoCollection(bestMatchedLogo).name
    end
end
```



```

score = matchPointNumber(1,1);
end
end

```

3.3 Reading Logo

After voting of the training logos, we decide the final winner. The name of that logo is played from a recorded voice file to the user. As a result, blind or visually impaired person can be now aware of the names of the surrounding shops.

A sample comparison between the training and the test frame is provided in the Figure 1. The image on the left is the training frame and the image on the right is the test frame. The feature points and their matching are shown in the figure as well.

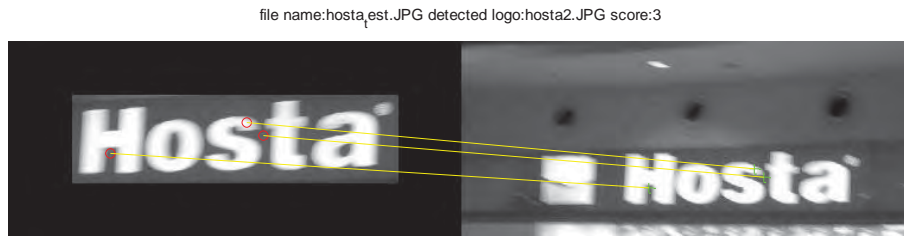


Figure 1. A sample comparison of the training and test images. The training image is given on the left, and test image is given on the right. For simplicity only three matching feature points are shown. The red circles show the feature points of the training image, and green + signs show the feature points of the test image. The yellow lines show the matching feature points.

4. PERFORMANCE TESTS AND RESULTS

We have collected 5 logo videos from a shopping mall with an average duration of 15 seconds. We have prepared Training Logo Database including 2 samples from each logo which makes 10 samples on total. To test the success of the proposed algorithm, we first extract all the frames including any logo images. As seen in Table 3, we have totally 69 test frames for 5 different logos. Using only a training database with 10 sample logos, we classify these 69 test frames as described above.

The results in Table 3 are given in percentage of the successfully recognized logo frames to the total test frame numbers. We further report the ratio of misclassification incidents and undetected cases. For example, the logo of Hosta company has a test frame set of size 8, and 5 of them are correctly classified, however only 2 of them are misclassified to be appeared as different logos. Furthermore, 1 of the test logo is undetected. In the discussion section, we present some justifications about the reported results considering the given videos and Training Logo Database.

Table 3. Confusion matrix of the classification performance of the proposed method.

| | Hosta | Otantik | Özsüt | Taşfırın | Teppanyaki | Undetected | Total | Success (%) |
|--------------|-------|---------|-------|----------|------------|------------|-----------|-------------|
| Hosta | 5 | 1 | | | 1 | 1 | 8 | 62.5 |
| Otantik | | 8 | | | | 1 | 9 | 88.9 |
| Özsüt | | | 8 | | | 1 | 9 | 88.9 |
| Taşfırın | | | | 13 | 1 | - | 14 | 93.6 |
| Teppanyaki | | | | | 29 | - | 29 | 100 |
| Total | | | | | | | 69 | 86.7 |

As can be seen from Table 1, some of the logos have higher number of test images than the others. For example Teppyenyaki has a total of 29 test images, and Hosta has a total of only 8 test images, etc. This is because of the video capturing speed of the users. When the video capturing speed is low, the blurring effect of the camera is also low, which also increase the performance of the classification.

It can be seen that Hosta company has the lowest classification performance when compared with the other companies. This can be due to the fact that Hosta has a classical logo style (as can be seen in Figure 1) and the feature points of the Hosta may be confused with the other logos.

5. CONCLUSIONS

As a conclusion, we can argue that the proposed algorithm is able to recognize the logos in the given videos with an high success ratio as much as 87%. Even the videos are recorded from different angles and distance with respect to the logos and under different lighting conditions and noise levels, the proposed algorithm manages to determine the logos embedded in the videos successfully. Secondly, we observe that the computation takes considerably less time such that the results could be reported near real time.

On the other hand the work needs further improvements. First of all the logo database should be extended to include more logos of the companies. After extending the logo database performance tests should be renewed, since extending the training dataset will extend the processing power and processing time. Besides similar logo types may be confusing and because of that performance may degrade.

Another important point about the results is that the video capturing speed may directly influence the classification performance of the proposed system. When the videos are captured very fast, then the cameras introduce motion blur effect. The motion blur effect may influence the feature point detection performance, which directly degrades the classification performance. But on the other hand motion blur effect may be reduced by using sharpening filters, which is not considered in this work.

Acknowledgements: This work is supported by Atılım University by the project number ATU-LAP-C-1314-07.

6. REFERENCES

- [1] Merriam-Webster Dictionary [Online].
Available: <http://www.merriam-webster.com/dictionary/logo>
- [2] Psyllos, Apostolos P., C-NE Anagnostopoulos, and Eleftherios Kayafas. "Vehicle logo recognition using a SIFT-based enhanced matching scheme." *Intelligent Transportation Systems, IEEE Transactions on* 11.2 (2010): 322-328.
- [3] Li, Zhe, Matthias Schulte-Austum, and Martin Neschen. "Fast logo detection and recognition in document images." *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010.
- [4] Zhu, Guangyu, and David Doermann. "Logo matching for document image retrieval." *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*. IEEE, 2009.
- [5] Phan, Raymond, and Dimitrios Androustos. "Content-based retrieval of logo and trademarks in unconstrained color image databases using Color Edge Gradient Co-occurrence Histograms." *Computer Vision and Image Understanding* 114.1 (2010): 66-84.
- [6] Eldib, Mohamed Y., et al. "Soccer video summarization using enhanced logo detection." *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009.
- [7] Wang, Hongye, and Youbin Chen. "Logo detection in document images based on boundary extension of feature rectangles." *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*. IEEE, 2009.

- [8] Chen, Yu, and Vrizlynn LL Thing. "Using Low Level Gradient Channels for Computationally Efficient Object Detection and Its Application in Logo Detection." *Multimedia (ISM)*, 2012 IEEE International Symposium on. IEEE, 2012.
- [9] Sun, Lin, et al. "Application of affine invariant SIFT matching for automatic logo recognition." *Computer Engineering and Design* 32.4 (2011): 1406-1410.
- [10] Zhang, Yuan, et al. "Spatial connected component pre-locating algorithm for rapid logo detection." *Acoustics, Speech and Signal Processing (ICASSP)*, 2012 IEEE International Conference on. IEEE, 2012.
- [11] MATLAB, [Online]. Available: <http://www.mathworks.com/>
- [12] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding (CVIU)*, Vol. 110, No. 3, pp. 346–359, 2008.
- [13] Silpa-Anan, Chanop, and Richard Hartley. "Optimised KD-trees for fast image descriptor matching." *Computer Vision and Pattern Recognition*, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008.

Efficient Data Gathering in WSN with a Range Constrained Mobile Relay

I Cereci¹, H Dağlayan², N Kılınc³, S Aktaş⁴, M Karakaya⁵

¹Department of Computer Engineering, Atılım University,

İncek, Ankara, Turkey

¹*ibrahim.cereci@atilim.edu.tr*, ²*hazan.daglayan@atilim.edu.tr*, ³*nergiz.kilinc@atilim.edu.tr*,
⁴*senem.aktas@atilim.edu.tr*, ⁵*murat.karakaya@atilim.edu.tr*

ABSTRACT

In Wireless Sensor Networks (WSN), Mobile Relay (MR) is used to collect data from the sensors dispersed in a region. Sensors sense the surrounding environment and record the measurements into their memory. MR has some range constraint which limits its traveling distance in the region where the sensors are located. In this paper, we proposed a Genetic Algorithm (GA) based method to guide MR in such a way that the amount of the data gathered from the sensors are maximized for a given range value. Since the amount of data collected in the memory of each sensor is dynamically increasing with individual sensory rate, this problem brings out different challenges compared to well-known problems such as, Traveling Salesman Problem (TSP) and Vehicle Routing Problem (VRP). We have adapted GA to solve this problem by modifying crossover and mutation functions. We implemented two different mutation algorithms and compared their effects. Furthermore, since the sensory data values dynamically change with time, we couldn't use one-point crossover as it is, so we modified one-point crossover function to transform it into a suitable one. We have simulated the proposed GA method and obtained promising results under different simulation and algorithm settings.

Keywords – Wireless Sensor Networks, Mobile Relay, Genetic Algorithm, scheduling

1. INTRODUCTION

Wireless Sensor Networks (WSN) and their challenges are studied in depth due to its popularity in many works such as [1][2][3][4]. There are many different data collection approaches implemented which can be classified into two main categories according to the sink mobility: Static Sink (SS) and Mobile Sink (MS). When SS is implemented all the sensors in the network should route their data to the SS by using some wireless communications methods. On the other hand, if MS is chosen, MS itself visits the sensors and collects data directly from each sensor. Furthermore, MS can be categorized according to their data gathering function. If a Sink is collecting and transferring the data to a remote central, it is a Mobile Sink [1]. However, in some cases mobile element visiting sensors is not a sink at all but it can collect data and transfer it to a SS. Thus, this kind of mobile element is called Mobile Relay (MR) [1].

In this work, we focus on data collection with a MR. We propose a method to produce a schedule for the MR to visit sensors such that it can collect maximum amount of data at the end of its tour. In the following sections we first define the problem in details and then introduce the basic concepts of Genetic Algorithm (GA) which our solution is based on. In Section 4, we provide the details of the method by explaining how GA is adapted to solve the defined problem. Simulation environment and test results are presented and discussed in Section 5. Finally, we provide conclusions and future work direction in the last section.

2. PROBLEM DEFINITION

We assume that sensors are located in a grid topology and their locations and initial amount of data stored are known. SS is located in the middle of the grid. There is a single MR to collect the sensory data. MR's trip should start from and end in the SS. Sensors have infinite memory so that there is no memory overflow. Sensors have a constant sensing rate. MR velocity is fixed and data transfer time between MR and sensors is negligible. MR can travel at most according to the given travel range.

The problem is to prepare a sensor visiting schedule for MR such that it can collect maximum amount of data from the sensors and transfer them to the SS at the end of the tour. We name this problem as *Maximum Data Gathering by a MR with Limited Range* (MDGMR/LR).

3. GENETIC ALGORITHM

Genetic algorithm is one of the popular evolutionary meta heuristics which is applied to many different engineering problems successfully [5][6][7][8]. Genetic algorithms (GAs) were invented by John Holland in the 1960s. Development continued by Holland and colleagues in 1970s. Their main goal was not design algorithms for specific problems, but to study how adaptation occur in the nature and import them into computer systems [9].

GAs are a family of computational models inspired by evolution. They hold the potential solution in a chromosome-like data structure and apply operations such as crossover and mutation in order to improve the solution. They are meta-heuristics which commonly considered as optimizers [10]. Implementation of GA begins with the creation of the initial population (can be random), after initial population is created, the individuals are selected using a selection algorithm to be parents in the crossover process. Crossover method takes partial solutions from each parent and merges them as new solutions. The result of a crossover can be one or more off-springs. Later, mutation operation is applied to the off-springs in order to apply little changes to the solution to escape possible similar local solutions. This cycle continues until the maximum number of generations is reached or no more improvement between generations is gained.

4. MAXIMIZING GATHERED DATA

We have developed a solution based on the GA for the MDGMR/LR problem. Details of the proposed solution are provided below.

2.1 Schedule Validation

When a sensor visiting schedule (a chromosome) is produced by the proposed GA method, it needs to be validated according to the given assumptions and the constraints. For example, the total tour length of the schedule (chromosome) must be less than or equal to the given range of the MR. Tours also should begin and end with the location of the Static Sink (SS) which is located on the middle of the grid. As another example, identical tours should be eliminated in the population as well. Our validation algorithm takes any tour and fits it with the definition of a possible solution. Schedule validation algorithm is given in Table 1.

2.2 Preparing Initial Population

The initial population for the GA is produced selecting sensors randomly [10]. As the number of sensors depends on the tour length each individual can have different number of sensors. Therefore, we validate each individual using the schedule validation algorithm discussed. Total number of individuals in the initial population is a GA parameter to the system. After schedule validation, we are ready to continue with the next GA phase.

Table 1. *The Schedule Validation pseudocode.*

```
Procedure scheduleValidate(Schedule){
if ( checkStartingPosition(currentPosition) != staticSinkPosition)
    newSchedule.put(staticSinkPosition)

currentRange = calculateRange(Schedule);
update(currentPosition);

while(currentRange + returnToSSTime <RANGE)
    newSchedule.put(currentPosition, iteration);
    update(currentPosition);
    iteration++;
    duplication= checkDuplication(newSchedule, Population);
    if (duplication)
        newSchedule= mutate(Schedule);

if ( checkEndingPosition(currentPosition) != staticSinkPosition)
    newSchedule.put(staticSinkPosition)
return newSchedule;
```

2.3 *Applying Crossover*

Whenever two pairs are selected to produce off-springs, we apply one-point crossover method [11]. Firstly, the selection method we use in order to select the parent is Tournament Selection (TS) [12]. In TS we select random individuals in the population according to the tournament size, and then we compare the fitness values of those individuals and select the fittest, as one of our parents. We apply the same procedure for the second parent also. Generally speaking, gene size in GA is fixed, but due to nature of our problem we have variable size genes. So we modified the one-point crossover algorithm for preventing illegal divisions, or end up with a gene that will not reach the maximum range in our problem. After applying the crossover and producing the off-springs, we validate these off-springs according to our validation algorithm explained in Table 1.

2.4 *Applying Mutation*

Each off-spring is undergone the mutation process. The mutation rate is a variable in our GA algorithm which determines whether a gene will be subject to mutation or not. Firstly we adopted swap mutation [12]. Swap mutation takes two sensors in a tour and swaps their locations. This method didn't perform well in our case, since the mutation is only limited to the sensors that only come from the parents after their crossover. All the other sensors that parents didn't visit was ignored. As a solution to that, we developed a mutation (Insertion), in which a sensor is can be replaced with another sensor that is in the entire list of sensors of the grid topology. As a result of this mutation, an off-spring can have a sensor in its tour that neither of its parents visited before. This change, increased the convergence to better fitness values rapidly. When the mutation is finalized the mutated off-springs are validated.

2.5 *Selection of Candidates for Next Generation*

In this phase, we select new members of the next population among the parents and the off-springs according to the fitness value. We can decide the portion of the parent population that will be put to the calculation using a variable in the GA program. At the end, all individuals are sorted

according to their fitness values descending, and the ones that are at the bottom, whether they are parent or offspring are removed. Remaining individuals constitute the new generation.

2.6 Determining Fitness Value

Fitness value is defined as the amount of data that is gathered at the end of the given schedule. Applying crossover and mutation highly effect the fitness value, so it is recalculated each time the validation algorithm runs.

5. SIMULATION AND TEST RESULTS

We have implemented the proposed solution with Java language. The simulation parameters are summarized in Table 2. Simulation code is implemented according to the assumptions presented in Section 2 (Problem Definition), and the methods we use for each step of the algorithm are described in Section 4. The GA parameters are given in Table 3.

Table 2. *The Simulation Parameters.*

| Parameter | Default Value | Range | Notes |
|------------------------|---------------|-------------|---|
| Sensor Topology | Grid | - | 25 x 25 (625 total) sensors |
| Distance | 20 meters | - | Horizontal/Vertical Distance between two sensors |
| MR velocity | 40 k/hr | 30-50k/hr | Fixed |
| Sensing Rate | 1 KB/sec | 0.5-2KB/sec | Fixed sensing rate |
| Sensor Memory Capacity | Infinitive | | Memory of each sensor is initialized with a random amount of data |

Table 3. *The GA Parameters.*

| Parameter | Default Value | Range | Notes |
|-------------------|--------------------------|------------|-------|
| Population Size | 200 | 100 - 400 | |
| Generation Number | 1000 | 500-2000 | |
| Crossover Method | Modified One-Point | | |
| Mutation Method | Swap Position, Insertion | | |
| Mutation Rate | 0.005 | 0.0005-0.1 | |

We have conducted a set of experiments. In the first experiment set, we observed the success of the resulting schedule. In the second experiment, we changed the GA Population Size and Generation Number parameters to monitor their effects on the convergence of the proposed method. In the third experiment, we modified the speed of the MR to observe the speed effect on the data collection. As a last experiment, we used different sensing rates for the sensors and record its impact on the target function. As discussed in details below the proposed method has adapted all these different parameter settings positively by producing successful schedules. All the runs are done 20 times, averages and standard deviation are calculated accordingly.

3.1 Base Experiment

In the first experiment we apply the default values of the population and generation size values in order to come up with an appropriate base experiment. In Figure 1, we explored the different generation sizes varying from 100 to 4000 while keeping population size in 200. From the figure we can see that after 1000th generation the fitness starts to slow down and after 1500th generation stays stable. It can also be observed that before 500th generation there is a dramatic drop in the fitness. So we chose our base generation to be 1000, ranging from 500-2000. In Figure 2, one can see the effect of increasing the population size. It can be observed that between population 50 and 100 there is a much bigger gap in fitness than populations between 200 and 800. So increasing the

population after certain value does not speed up the convergence drastically. For our experiments we keep the mutation rate as 0.005, since it was the fastest (most of the time the only) converging rate to solution.

As a result of the above observations, we set the population size and the generation number 200 and 1000 respectively.

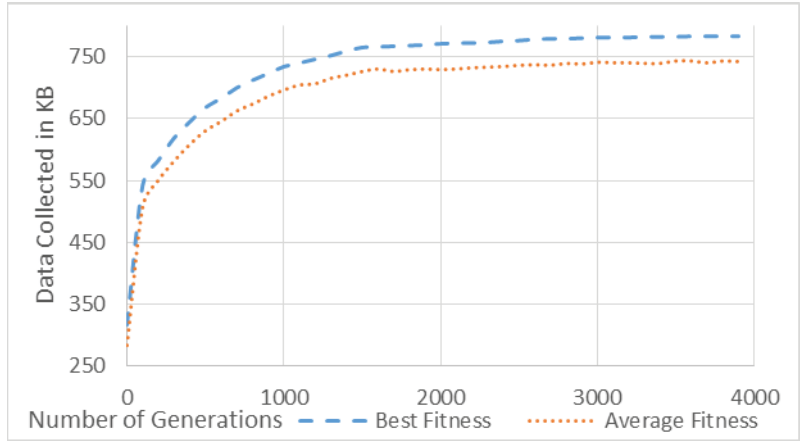


Figure 1. Increase in the Fitness Value due to the number of generations

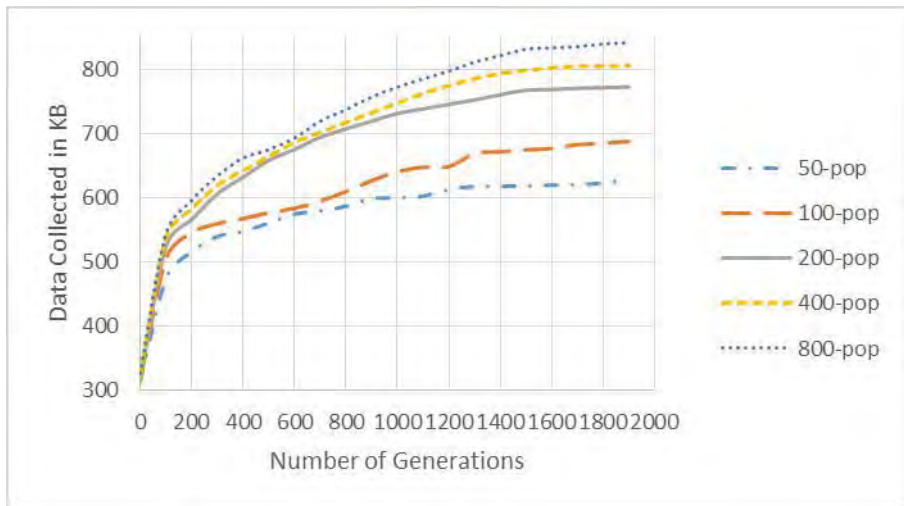


Figure 2. Effects of different population sizes varying from 50 to 800 on the fitness value.

3.2 Changing GA Parameter Values

We have designed three setups at the boundaries of our experiment range: Low, Standard, and High. At Low setup, we choose 100 as population size, 500 as generation size. In the Standard setup, population is 200 and the generation is 1000. Lastly at High setup, population is 400 and generation is 2000. We can see the comparison of their fitness values with their standard deviation in Figure 3. From the figure we can see that increasing the population size and generation size steadily increases the fitness values. However, the run time also increase dramatically from the standard setup to the high. A standard setup results with an answer, under 30 seconds, but a high setup problem takes around 10 minutes to be solved on an Intel I7 3.7Ghz, 16GB RAM machine. One can say that, standard setup gives an acceptable result in an acceptable time. As for mutation, when Swap-Position algorithm is selected in the standard GA setup, we only collected 657KB data, where with Insertion, we collected 737KB. Making Insertion a better choice for our case.

Table 4 shows us the effect of changing the mutation rate on the best fitness value. It can be seen that the value 0.005 should be chosen as default in order to maximize the fitness.

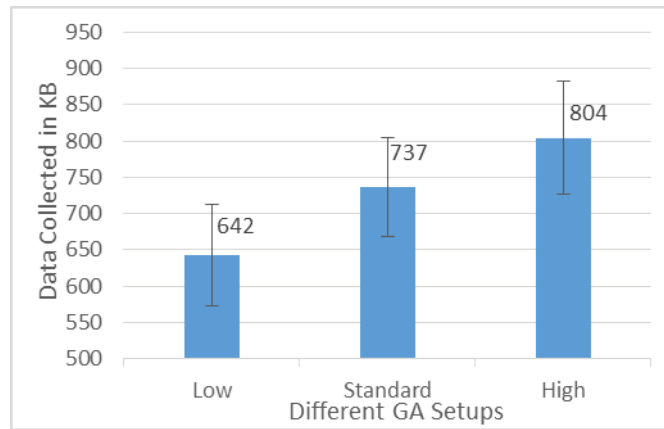


Figure 3. The difference in Fitness between Low, Standard and High GA settings.

Table 4. The effect of Mutation Rate on Fitness.

| Mutation Rate | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1 |
|------------------------|--------|-------|-------|------|------|-----|
| Data Collected (in KB) | 646 | 680 | 752 | 719 | 658 | 599 |

3.4 Different MR Speeds

The effects of different MR speeds can be observed from Figure 4. It is seen that decreasing the MR speed from 40Km/h to 30Km/h has a bigger impact than increasing speed from 40Km/h to 50Km/h. This is caused by MR traveling through the grid environment and collecting data from the sensors. Although we go faster, if the number of revisits to the sensors increase, impact of the speed increase will be less.

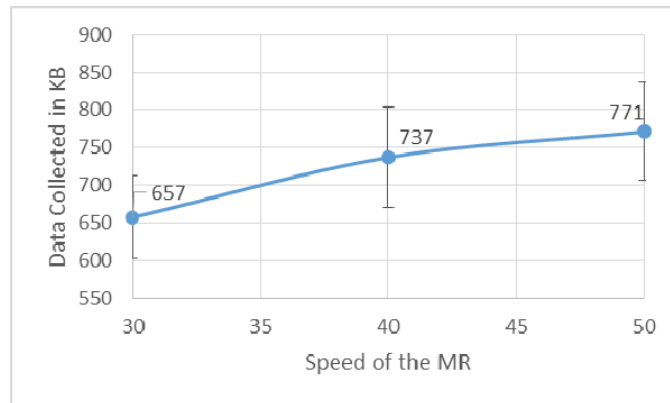


Figure 4. The effect of the Mobile Relay speed on the fitness value

3.5 Different Sensing Rates

As the last experiment, we changed the sensing rates of the sensors between 0.5 and 2, in order to see the impact. Figure 5, depicts that increasing the sensing rate improves the fitness value steadily. Since the sensing rate increases the overall data values in each sensor, it is expected to affect the result steadily.

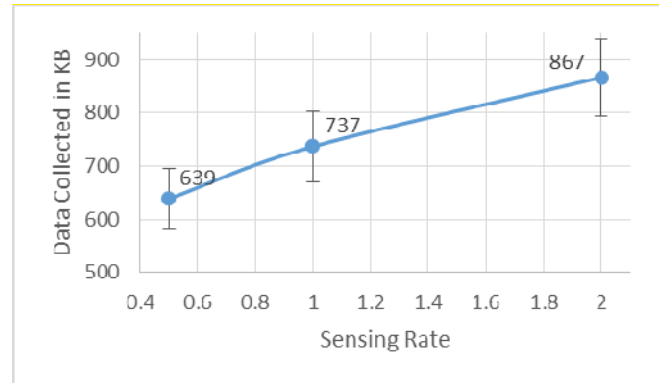


Figure 5. *The effect of sensing rate of the sensors on the fitness values.*

5. CONCLUSIONS

In this work we focus on the data gathering in WSN using a MR with a limited range. We have adapted the GA to solve this problem and implemented the solution. We have conducted various experiments to observe the impacts of the simulation and GA parameters on the behavior of the proposed methods. The results show that under different simulation settings, the proposed method produces efficient schedules to collect maximum data.

As a future work we aim at designing a novel method for scheduling more than one MR to gather data for the given MDGMR/LR problem. We also aim to include a heuristic method in our initial population generation and mutation, in order to converge to better fitness values.

6. REFERENCES

- [1] Di Francesco, Mario, Sajal K. Das, and Giuseppe Anastasi. "Data collection in wireless sensor networks with mobile elements: A survey." *ACM Transactions on Sensor Networks (TOSN)* 8.1 (2011): 7.
- [2] Akyildiz, Ian F., et al. "A survey on sensor networks." *Communications magazine, IEEE* 40.8 (2002): 102-114.
- [3] Anastasi, Giuseppe, et al. "Energy conservation in wireless sensor networks: A survey." *Ad Hoc Networks* 7.3 (2009): 537-568.
- [4] Karakaya, Murat. "Deadline-aware energy-efficient query scheduling in wireless sensor networks with mobile sink." *The Scientific World Journal* 2013 (2013).
- [5] Zhou, Aimin, et al. "Multiobjective evolutionary algorithms: A survey of the state of the art." *Swarm and Evolutionary Computation* 1.1 (2011): 32-49.
- [6] Savuran, Halil, Karakaya, Murat. "Route Optimization Methods for Unmanned Air Vehicle Launched from a Carrier", 4th International Conference on Computer and Software Modeling (ICCSM 2014), Bangkok, Thailand, 2014.
- [7] Sevinc, Ender, Karakaya, Murat. "Maximizing UAV Target Coverage under Flight Range and Target Service Time Constraints", 4th International Conference on Computer and Software Modeling (ICCSM 2014), Bangkok, Thailand, 2014.
- [8] Guo, Pengfei, Xuezhong Wang, and Yingshi Han. "The enhanced genetic algorithms for the optimization design." *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*. Vol. 7. IEEE, 2010.
- [9] Mitchell, Melanie. "An Introduction to Genetic Algorithms (Complex Adaptive Systems)" Published by: MIT press, Feb 6, 1998
- [10] Whitley, Darrell. "A genetic algorithm tutorial." *Statistics and computing* 4.2 (1994): 65-85.
- [11] Poli, Riccardo. "Exact Schema Theory for Genetic Programming and Variable-Length Genetic Algorithms with One-Point Crossover", *Genetic Programming and Evolvable Machines*, pg. 123-163, 2001
- [12] Miller, Brad L., and David E. Goldberg. "Genetic algorithms, tournament selection, and the effects of noise." *Complex Systems* 9.3 (1995): 193-212.

Differential Search Algorithm with Levy Flight

Emre Sertel¹, Oguz Altun²

Abstract

In solution of optimization problems, metaheuristic algorithms inspired from nature widely took place. These algorithms (e. g. Cuckoo Search, Bat, and Artificial Bee Colony) imitate movement behaviour of superorganisms to obtain optimum solution. One of recent naturally inspired algorithm, differential search algorithm has got successful results on many common test problems. In this algorithm, a type of random walk, Brownian motion is applied to simulate migration of superorganism which looks for food.

In this paper, instead of Brownian motion another random walk model, Levy flight is integrated with differential search algorithm to gain angle to search path. New algorithm is compared to the original DS algorithm by testing on optimization test problems, rastrigin, rosenbrock and sphere.

1. Introduction

Most classical algorithms are deterministic. And those algorithms work well on problems which have smooth distinct distribution. If the problem's objective functions' values are distributed discontinuously, classic algorithms do not work well. It is needed non-gradient algorithms, stochastic algorithms to find optimum result. There are two types of stochastic algorithms, heuristic and metaheuristic. Even it can be classified that metaheuristic is subtopic of heuristic because their difference is very small. Heuristic algorithms struggle to approach to the optimum solution among possible solutions in an objective function. A new stochastic algorithm type, metaheuristic firstly introduced in Glover in 1986, brings together basic heuristic technics in higher lever structures and searching efficiently. [1]

¹ emresertel@gmail.com, Epoka University, Computer Engineering Department, Tirana, Albania

² oguz@ce.yildiz.edu.tr, Yildiz Technical University, Faculty of Electrical and Electronics Engineering, Computer Engineering Department, Istanbul, Turkey

Metaheuristic can be simply defined as complete strategies, methods for search process. It actually deals with optimizing problems to find effective solutions as possible. However, methods of metaheuristic never promise a perfect solution. Instead, its approach tries to reach more efficient solution to corresponding problem. [2] It means solutions that metaheuristic algorithms found are approximate results instead of deterministic. This originates from nature of optimization problems. For many years, different algorithms have been developed like artificial bee algorithm, hill climbing algorithm, particle swarm optimization algorithm, cuckoo search algorithm. One of most recent algorithms is differential search algorithm developed by P. Civicioglu.

Metaheuristic doesn't guarantee to return best solution, but solution it produced in a short period is good enough to implement. Metaheuristic algorithms are not created for specific problems. Consequently, they can be applied to a wide range of optimization problems. However, different approaches can obtain better results on different problems.

This paper is organized as follows. In Section 2, literature review is done. Classification of metaheuristics, nature-inspired metaheuristics, Brownian motion and Levy Flight are discussed to make work clear to understand. In Section 3, test functions are introduced. In Section 4, Differential Search Algorithm is analysed. In Section 5, Differential Search (DS) Algorithm with Levy Flight is introduced. In Section 6, DS and new algorithm are compared.

2. Literature Review

A. Classifications of Metaheuristics

There are different viewpoints to classify metaheuristic algorithms contingent on characteristic methods that they carry out to explore search space. [3] It is taken into consideration algorithm of this proposed paper to mention classification ways of metaheuristics. Therefore, metaheuristic methods are discussed according to their origin and number of solutions at a time.

I. Population based and single point search

Amount of candidate solutions implemented at a time, is one of most important characteristics (distinction) to categorize metaheuristics. In the single point search, only one candidate solution is produced at each iteration of the optimization algorithm. These kinds of algorithms are called trajectory methods. Simulated annealing, Tabu search and variable neighbourhood search are such single point method algorithms. [4] They include metaheuristics based on local search. On the other hand, population based

algorithms carry out search steps which marks progress of candidate solutions in the search space. In each iteration, population is modified using pheromone trails like in ant colony, artificial bee and cuckoo search optimization algorithms. [4]

II. Nature-inspired and non-nature inspired

There are non-nature inspired algorithms such as Tabu search, iterated local search and grasp. However, many metaheuristic methods are really inspired by nature phenomena. Their intention mainly is imitating movement of organisms and biological mechanisms. Ant colony algorithm tries to simulate behaviour of ant colony seeking food source. The Genetic algorithm, created by John Holland and his colleagues, is a simulation of natural selection based on Charles Darwin's theory in biology. Crossover, recombination, mutation and selection are implemented to algorithm as a search strategy to solve optimization problems. [5] Another nature-inspired algorithm, firefly algorithm, formulated by Xin-She Yang at Cambridge University in 2007, is inspired by fireflies which have rhythmic flashing light using to attract other fireflies.

B. Evolutionary Algorithms

Evolutionary algorithms are inspired by evolution of living beings to adapt well to the environment. It uses evolution operators called recombination, mutation and selection as steps of strategy to find an optimum solution in search space. These algorithms are population based and nature-inspired metaheuristics. Better solutions (fitness) survive on a set of candidate solutions to approach gradually to the optimum. Therefore, a new population is created by selecting solutions which have better results on objective function after each iteration until algorithm converges to best solution in a reasonable period. [6]

General steps of evolutionary algorithms:

1. Initialization of population
2. Selections
3. Generating new candidate solutions (Recombination)
4. Evaluating candidate solutions in the population
5. Repeating from step 2 until convergence criteria is satisfied.

C. Metaheuristic inspired from Nature

In optimization problems, where solution can be found is not known. And it is nonsense to check every single corner to find the best solution in large search space within a time limit. Therefore, what is done in metaheuristic methods is trying to discover optimum solution by doing some random walk while searching for a clue which can take solution to better fitness. This kind of random movement is main characteristic of metaheuristic. Thus, each time search algorithm is run, another optimal solution is found but solutions must be close to each other for a good algorithm. The aim in metaheuristic is to create a simpler algorithm that works efficiently and finds good solutions. [1]

One of methods to create such random walks is imitating movement of foraging animals, superorganisms. Many different species of predatory superorganisms check fertility of the site they want to migrate. If the site they checked has adequate food that they need, they settle in the new site for a while. Then they continue their migration to another site to find more fertile areas. [7]

There are many algorithms that model the movement of superorganisms (e.g., Cuckoo Search, Artificial Bee Colony, Particle Swarm Optimization, Ant Colony, and Differential Search).

D. Random Walk

A random walk is a mathematical formalization of a path that consists of a succession of random steps. For example, the path traced by a molecule as it travels in a liquid or a gas, the search path of a foraging animal can be modeled as random walks. [1]

Mathematically, let S_N denotes the sum of each consecutive random step X_i , then S_N forms a random walk

$$S_N = \sum_{i=1}^N X_i = X_1 + \dots + X_N \quad (1)$$

where X_i is a random step drawn from a random distribution. This relationship can also be written as a recursive formula

$$S_N = \sum_{i=1}^{N-1} X_i + X_N = S_{N-1} + X_N \quad (2)$$

which means the next state S_N will only depend the current existing state S_{N-1} and the motion or transition X_N from the existing state to the next state. [1]

E. Brownian motion

Brownian motion is a random movement of a particle resulting from colliding with molecules in a gas or liquids. It is a model to represent random movements mathematically. This mathematical model has used in different science branch to apply several real-world applications. [8]

F. Levy Flight

Levy flight is a kind of random walk which its step lengths are obtained randomly from a probability distribution with a power law tail. Levy flight distribution can be mathematically defined by formula (3).

$$L(s, \gamma, \mu) = \begin{cases} \sqrt{\frac{\gamma}{2\pi}} \exp\left[-\frac{\gamma}{2(s-\mu)}\right] \frac{1}{(s-\mu)^{3/2}}, & 0 < \mu < s < \infty \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where $\mu > 0$ is a minimum step and γ is a scale parameter.

This type of random walk has been found in nature. Many different foraging animals like sharks, albatrosses, fruit flies, penguins, tuna and turtles all do typical levy flight movement while they search for food. [9] [10]

Levy flight can improve the efficiency of resource searches in unknown areas. Levy flight has been used in many metaheuristic algorithms inspired from nature. For instance, Cuckoo search algorithm uses parasitic behaviour of some cuckoos with combination of levy flight motion of some birds and fruit flies. [11]

The chaotic Levy flight for the improved bat algorithm proposed by Jiann-Honng Lin also used levy motion to enrich search behaviour and avoid to be trapped in local optimum. [12] Another study formulates combination of levy flight and firefly algorithm.

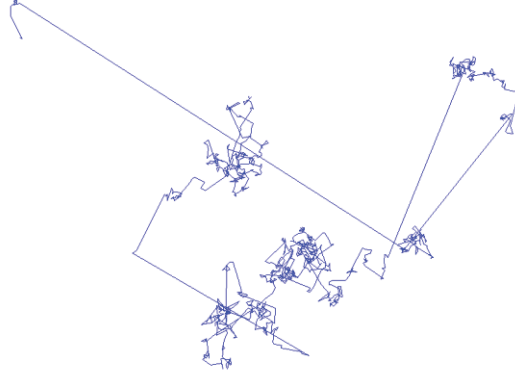


Figure 1 Typical Levy Flight plot with 1000 steps

In Mantegna's algorithm, accurate algorithm for numerical simulation of Levy stable stochastic processes, the step lengths s can be calculated by

$$s = \frac{u}{|v|^{1/\beta}}, \quad (4)$$

where v and u are drawn from normal distributions. That is

$$u \sim N(0, \sigma_u^2), \quad v \sim N(0, \sigma_v^2),$$

where,

$$\sigma_u = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \quad \sigma_v = 1. \quad (5)$$

Here $\Gamma(z)$ is the gamma function

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt. \quad (6)$$

3. Test Functions

To evaluate and compare algorithms, common optimization test problems are used negative rastrigin, negative sphere, and negative rosenbrock functions.

A. Negative Rastrigin Function

Rastrigin function which is first introduced by Rastrigin is a non-convex and a typical non-linear multimodal function. [2]

$$\begin{aligned} &\text{maximize} && f(\langle x_1, \dots, x_n \rangle) = -10n - \sum_{i=1}^n x_i^2 - 10 \cos(2\pi x_i) \\ &x_i \in [-5.12, 5.12] \end{aligned} \tag{7}$$

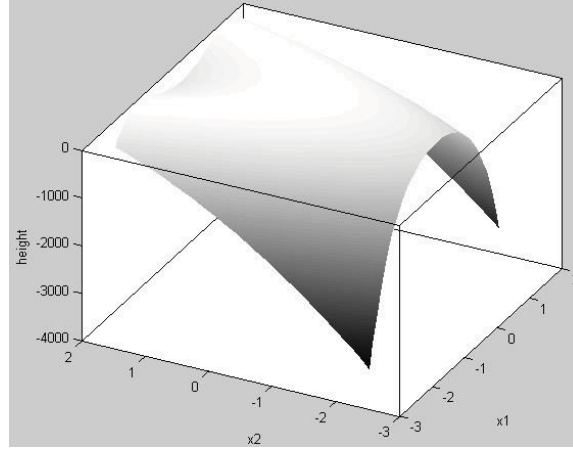


Figure 2 Plot of Negative Rastrigin Function

B. Negative Rosenbrock Function

Rosenbrock function is a non-convex function and introduced by Howard H. Rosenbrock in 1960. [2]

$$\begin{aligned} &\text{Maximize} && f(\langle x_1, \dots, x_n \rangle) = - \sum_{i=1}^{n-1} (1 - x_i)^2 + 100(x_{i+1} - x_i^2)^2 \\ &x_i \in [-2.048, 2.048] \end{aligned} \tag{8}$$

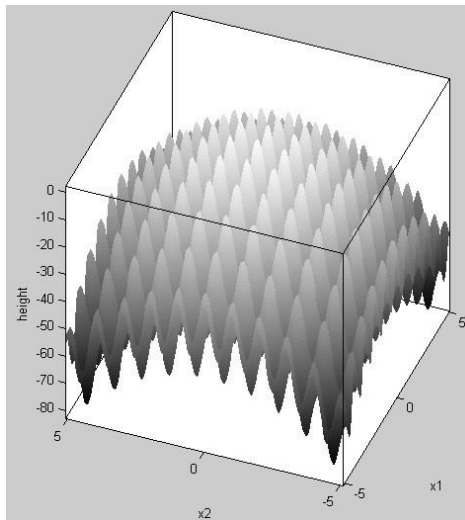


Figure 3 Plot of Negative Rosenbrock Function

C. Negative Sphere Function

Sphere function is a convex, unimodal and continues function. It is one of simplest function and part of a De Jong's test suite so it is also called first function of De Jong. [2]

$$\text{Maximize } f((x_1, \dots, x_n)) = - \sum_{i=1}^n x_i^2 \quad x_i \in [-5.12, 5.12] \quad (9)$$

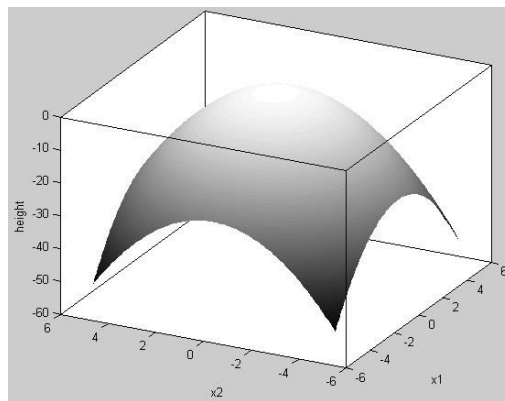


Figure 4 Plot of Negative Sphere Function

4. Analysis of “Differential Search Algorithm”

Differential search (DS) algorithm is built to solve optimization problems. DS algorithm imitates the Brownian motion of an organism to migrate. [7]

Amount of food and its sufficiency in different areas change in the nature because of the climatic changes through the year. Therefore, many different species of living creatures (e. g. many species of birds) migrates to a proper habitat where they can find food and water during season changes. Such as many species of birds migrates south for the winter, some insects like species of beetles, butterflies and moths have seasonal movement. And those migration movements have large numbers of individuals, superorganism. Migration of a superorganism can be simulated by using a type of random walk, Brownian motion. [7]

In Differential search algorithm, population is generated randomly as a candidate solution of the corresponding problem. In this algorithm artificial superorganism moves to global minimum of the problem so it optimizes fitness to minimum. While superorganism migrates, it checks whether the positions randomly selected are proper to settle for a short period. When the new position supplies adequate food, the member of the superorganism

temporarily stays in the stopover site during the migration. Later on, it goes on the migration from this new position. [7]

The new position that member of superorganism moves is calculated by getting distance to another randomly selected member and multiplying this distance by a scale value generated using a gamma random number generator in a range [0 1]. And the member takes this path to move towards donor. The scale allows changing position respectively to another member in the habitat. Stopover site is produced by using equation 11. [7]

$$\text{StopoverSite} = \text{Superorganism} + \text{Scale} * (\text{donor} - \text{Superorganism}) \quad (11)$$

Pseudo Code of Differential Search Algorithm

1. Initialize superorganism positions.
2. Get the artificial organism which has best food site. (y_1 , best fitness).
3. While $\text{assessment_count} \leq \text{assessments_number}$
 - a. Generate a donor by using randperm function.
 - b. Calculate scale.
 - c. Calculate stopover site with Brownian motion by getting difference of donor and superorganism.
 - d. Determine the directions for artificial organisms to immigrate
 - e. Define the new site in the determined directions.
 - f. If new site more fertilized
 - i. Organisms move to defined new site
4. Get the artificial organism which is in the new best food site (y_i , best fitness).
5. If $y_i > y_1$
 - a. Assign new best food site ($y_1 = y_i$)
6. Return best food site (best fitness)

5. Differential Search Algorithm integrated with Levy Flight

In this paper, two new algorithms with levy flight are proposed. In these algorithms, to simulate levy flight movement following equations are used to calculate stopover site.

In Mantegna's algorithm, the step lengths can be calculated by

$$step = \frac{u}{|v|^{1/\beta}}, \quad (12)$$

where v and u are pseudorandom values drawn from the standard normal distribution. That is

$$u = \text{randn}(2) * \sigma_u, \quad (13)$$

$$v = \text{randn}(2), \quad (14)$$

where

$$\sigma_u = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta} \quad (5)$$

$$\text{stepsize} = 0.01 * \text{step} * (\text{superorganism-donor}). \quad (15)$$

Here the factor 0.01 comes from the fact that $L/100$ should be the typical step size of walks/flights where L is the typical length scale. Otherwise, Levy flights may become too aggressive / efficient. [11]

$$\text{Stopover} = \text{superorganism} + \text{stepsize} * \text{randn}(2) \quad (16)$$

In first algorithm, beta coefficient stayed constant in 0.6. In second algorithm, beta value initialized to 0.6 but this time when better fitness is obtained in iteration, the value of beta is decreased by 0.01. When worse fitness is obtained in iteration, the value of beta is increased by 0.01. Thus, we do a shorter jump to a new position, after it is obtained better fitness.

Besides differential search and new algorithms have different random walk type, their migration direction is also different. While artificial organisms move towards their donor to a stopover site in differential search algorithm, artificial organisms in new algorithms change their direction from donor.

Pseudo code of DS with Levy Flight (constant alpha – DSL_C)

1. Initialize superorganism positions and $\alpha = 0.6$.

2. Get the artificial organism which has best food site. (y_1 , best fitness).
3. While assessment_count \leq assessments_number
4. Generate a donor by using randperm function.
5. Calculate stopover site with Levy flight by getting difference of donor and superorganism.
6. If stopover site more fertilized
7. Organisms move to stopover site
8. Get the artificial organism which is in the new best food site (y_i , best fitness).
9. If $y_i > y_1$
10. Assign new best food site ($y_1 = y_i$)
11. Return best food site (best fitness)

Pseudo code of DS with Levy Flight (variable alpha – DSL_V)

1. Initialize superorganism positions and $\alpha = 0.6$.
2. Get the artificial organism which has best food site. (y_1 , best fitness).
3. While assessment_count \leq assessments_number
 - a. Generate a donor by using randperm function.
 - b. Calculate stopover site with Levy flight by getting difference of donor and superorganism.
 - c. If stopover site more fertilized
 - i. Organisms move to stopover site
4. Get the artificial organism which is in the new best food site (y_i , best fitness).
5. If $y_i > y_1$

- a. Assign new best food site ($y_1 = y_i$)
 - b. Decrease α by $\alpha*0.01$
6. Else
- a. Increase α by $\alpha*0.01$
7. Return best food site (best fitness)

6. Comparison of Algorithms

Differential Search and proposed algorithms are run with 1000 assessments and 30 trials to solve maximization optimization problems (negative rastrigin, negative rosenbrock, and negative sphere) that are introduced in literature review chapter. These problems are negated to convert them to maximization optimization problem.

In order to compare performances of algorithms, the mean convergence graphs of assessments are plotted. On each test problems, convergence graphs show that new algorithms with levy flight reach better fitness faster than differential search algorithm.

A. Negative Rastrigin Problem

Table 1 shows that DS reaches higher fitness. Even if all three algorithms have successful results after 1000 assessments, DS obtains better fitness.

Table 1 Comparison of Results with Multi Runs on the Negative Rastrigin

| Algorithm | Maximum height | Mean height | Standard deviation of heights |
|-----------|----------------|-------------|-------------------------------|
| DS | -0.00048 | -0.55705 | 0.49108 |
| DSL_V | -0.02061 | -0.94563 | 0.69281 |
| DSL_C | -0.01129 | -1.0249 | 0.66427 |

When convergence graph in figure 5 is examined, it is seen that DSL_C and DSL_V algorithms are converged faster than DS in early assessments. But DS gets better heights after 350 assessments.

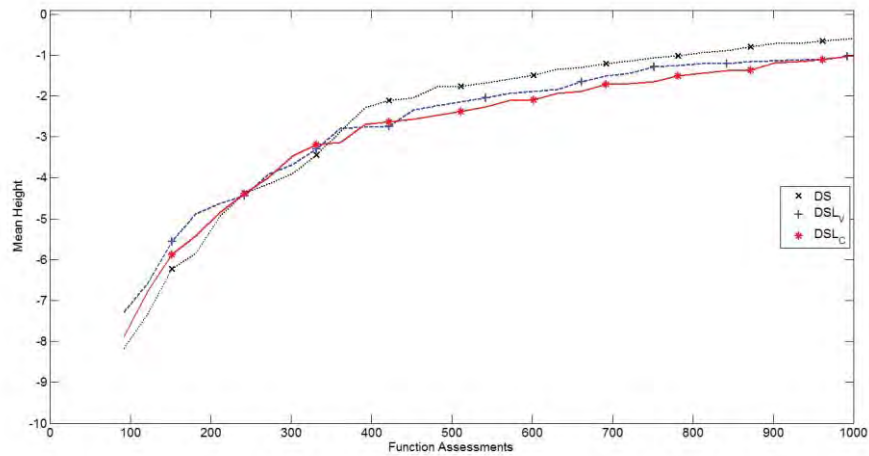


Figure 5 Convergence graph of algorithms on Negative Rastrigin problem

B. Negative Rosenbrock Problem

In negative rosenbrock problem, DS has better mean height and gets maximum height among all algorithms. This reported in table 2.

Table 2 Comparison of Results with Multi Runs on the Negative Rosenbrock

| Algorithm | Maximum height | Mean height | Standard deviation of heights |
|-----------|----------------|-------------|-------------------------------|
| DS | -1.12E-05 | -0.01656 | 0.016811 |
| DSL_V | -0.00022 | -0.07517 | 0.085616 |
| DSL_C | -0.00165 | -0.06035 | 0.055599 |

Again, DSL_C and DSL_V algorithms converge earlier than DS algorithm, when we check figure 6. On the other hand, all three algorithms obtain satisfying results after 500 assessments. However, DS obtains better after 200 assessments.

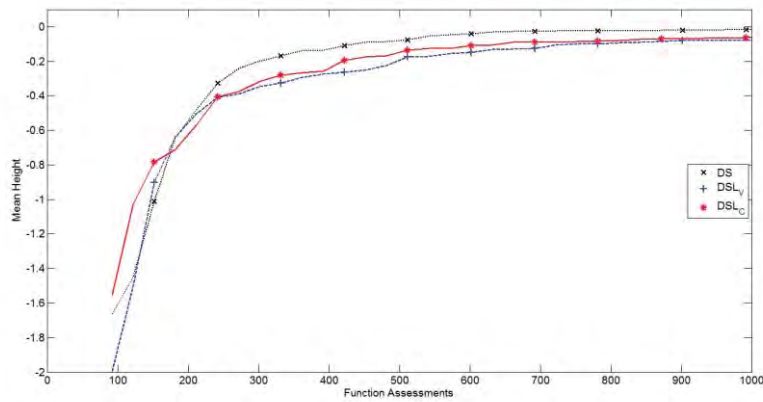


Figure 6 Convergence graph of algorithms on Negative Rosenbrock problem

C. Negative Sphere Problem

In negative sphere problem, all three algorithms reach higher positions before 300th assessment. This can be clearly seen in figure 3. This early success is resulted from shape of sphere function.

Table 3 Comparison of Results with Multi Runs on the Negative Sphere

| Algorithm | Maximum height | Mean height | Standard deviation of heights |
|-----------|----------------|-------------|-------------------------------|
| DS | -1.68E-06 | -0.00087 | 0.002153 |
| DSL_V | -8.48E-06 | -0.00048 | 0.000502 |
| DSL_C | -4.58E-06 | -0.00226 | 0.0074 |

However, new algorithm, DSL_V obtains better mean height and standard deviation of heights.

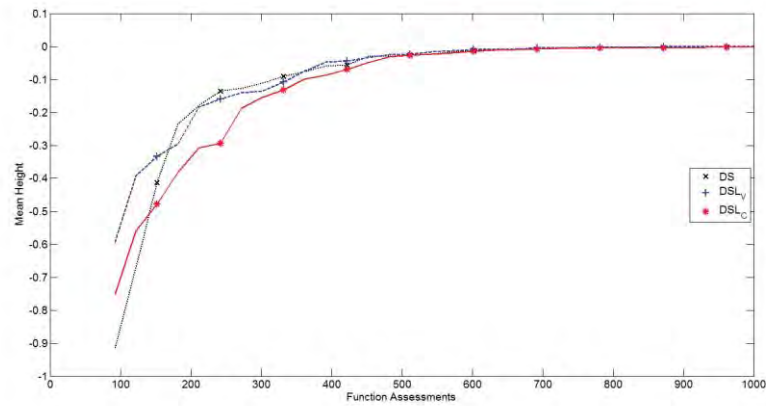


Figure 7 Convergence graph of algorithms on Negative Sphere problem

D. Conclusions

In this proposed research, one of latest metaheuristic algorithm, Differential Search is analysed in details. Random walk types, Brownian motion and Levy Flight are examined. It is applied Levy Flight to Differential Search instead of Brownian motion. DS and algorithms integrated with Levy Flight are compared on common objective functions, negative rastrigin function, negative rosenbrock function and negative sphere function.

Convergence graphs of algorithms, DS, DSL_V and DSL_C on each objective function is compared. All examined graphs shows that DS algorithm obtains better results than proposed algorithms. Therefore, DS algorithm with Brownian motion gets better fitness than DS with Levy Flight.

Bibliography

- [1] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*, Frome, England: Luniver Press, 2010.
- [2] S. Luke, *Essentials of Metaheuristics*, Raleigh, North Carolina: Lulu Press, 2009.
- [3] M. Birattari, L. Paquete, T. Strutzle and K. Varrenttrapp, "Classification of Metaheuristics and Design of Experiments for the Analysis of Components," AIDA-01-05, Darmstadt, Germany, 2001.
- [4] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Computing Surveys (CSUR)*, vol. 35, no. 3, pp. 268-308, 2003.
- [5] T. Weise, "Global Optimization Algorithms - Theory and Application," 26 06 2009. [Online]. Available: <http://www.it-weise.de/projects/book.pdf>. [Accessed 10 02 2013].
- [6] N. Kokash, "An introduction to heuristic algorithm," Department of Informatics and Telecommunications, Trento, Italy, 2005.
- [7] P. Civicioglu, "Transforming geocentric cartesian coordinates to geodetic coordinates by using differential search algorithm," *Elsevier*, pp. 229-247, 2012.
- [8] P. Morters and Y. Peres, *Brownian Motion*, Cambridge, England: Cambridge University Press, 2008.
- [9] G. M. Viswanathan, V. Afanasyev and S. V. Buldjrev, "Levy Flight Search Patterns of Wandering Albatrosses," *Nature*, vol. 381, no. 30, pp. 413-415, 1996.
- [10] D. W. Sims, N. E. Humphries, R. W. Bradford and B. D. Bruce, "Lévy flight and Brownian search patterns of a free-ranging predator reflect different prey field characteristics," *Animal Ecology*, vol. 81, no. 2, pp. 432-442, 2012.
- [11] X.-S. Yang and S. Deb, "Cuckoo Search via Levy Flights," *Proc. of World Congress on Nature & Biologically Inspired Computing*, pp. 210-214, 2009.
- [12] J.-H. Lin, C.-W. Chou, C.-H. Yang and H.-L. Tsai, "A Chaotic Levy Flight Bat Algorithm for Parameter Estimation in Nonlinear Dynamic Biological Systems," *Journal of Computer and Information Technology*, vol. 2, no. 2, pp. 56-63, 2012.

Evaluation of Semantic Similarity Measurement Algorithms For Word Sense Disambiguation

Şeyma Altun¹, Elton Domnori²

¹Computer Engineering, Epoka University,
Tirana, ALBANIA

²Computer Engineering, Epoka University,
Tirana, ALBANIA

[¹seyma.altun@gmail.com](mailto:seyma.altun@gmail.com), [²edomnori@epoka.edu.al](mailto:edomnori@epoka.edu.al)

ABSTRACT

Many words have more than one meaning. These words are known as polysemious. Word Sense Disambiguation (WSD) is the process of identifying the correct sense of a polysemious word in a context.

WSD algorithms measure the similarity between two words with different methods. In this paper, Java WordNet Similarity (JWS) and DISCO are represented by giving the detailed information of their similarity measurement metrics.

The aim of this article is to evaluate the similarity metrics behind the research of WSD algorithms. The data used for the evaluation of the algorithm is taken from Keymantic test set[1]. There are two different datasets named as Unimore and IMDB. The former is a dataset containing keywords and elements related with university. The latter is related with the cinema. Each dataset consists of the sample keyword queries posed by some voluntaries and data elements and attributes expected. A similarity value is measured using the similarity metrics in the JWS and DISCO for each pairs of keyword and data element.

Keywords – semantic similarity; word sense disambiguation

1. INTRODUCTION

Human languages are ambiguous because some words have more than one meaning. These words are called polysemious. Word Sense Disambiguation (WSD) is the task of identifying the appropriate meaning for such words in a given context.

For example, according to the Merriam-Webster dictionary [2], the word *bar* has multiple meanings, some of which are shown in following sentences:

(a) *There were bars in the windows to prevent escape.*

(b) *He bought a hot dog and a coke at the bar.*

In the first sentence *bar* means “a rigid piece of metal or wood; usually used as a fastening or obstruction or weapon”. The second has the sense of “a counter where you can obtain food or drink”.

WSD is important for the natural language applications such as machine translation, question answering, text classification, text summarization, and information retrieval. There are several approaches for WSD.

In this article, two packages having different similarity metrics are used for measuring semantic similarity to disambiguate the words. These are WordNet::Similarity (JWS)[3] and DISCO[4]. We will evaluate these similarity metrics with our dataset. The dataset is taken from Keymantic[1] work.

The accuracies of the similarity metrics are compared using the *precision value*, which is the proportion of true matches to all matches.

2. SIMILARITY MEASUREMENT ALGORITHMS

Several methods are proposed for measuring semantic similarity between words. These methods can be grouped as knowledge-based, supervised and unsupervised. In this section, similarity measurement algorithms in two different packages will be introduced.

2.1 DISCO

Kolb [4] developed the *DISCO* package which retrieves the *distributional similarity* between two words and lists collocations of the words in a context. Distributional similarity represents the similarity of words that have more collocations in common. Collocations are the words that are used together in a context. DISCO uses different corpuses as data source. In this work the British National Corpus (BNC) and English Wikipedia to carry out our evaluation has been used.

The algorithm has a pre-processing step in which the corpus is tokenized and highly frequent words are eliminated. Context window that defines how many words to check, size is 3. The algorithm checks three left and three right words for counting word co-occurrences. It shows the distribution of a word in a way that shows the ordered pairs of word and window position [5].

Disco program calculates two different similarity values using this information. These are: first-order and second-order similarity.

- First-order similarity is computed based on the collocations of two words.
- Second-order similarity between two words is computed based on their sets of distributional similar words. If two words have more common collocations, these words are distributional similar.

2.2 JAVA WORDNET::SIMILARITY

David Hope [3] created a package that is the Java implementation of WordNet::Similarity Perl package [6]. WordNet::Similarity [6] measures the semantic similarity and relatedness between word senses using WordNet lexical database and WordNet Info-Content Files that consist of some information collected from several corpora like BNC, Semcor etc. WordNet is a lexical database that shows the relations between word senses. The most important relations are hypernym and hyponym. Hypernym gives a broad group of the word. Hyponym is the more specific word or phrase.

JWS uses ten different similarity and relatedness measurements. These are: Adapted Lesk, Adapted Lesk Tanimoto, Adapted Lesk Tanimoto No Hyponyms, Hirst-St.Onge, Jiang-Conrath, Leacock-Chodorow, Lin, Path, Resnik and Wu-Palmer. In the following, these measurements are presented.

Adapted Lesk, Michael Lesk [7] proposed a disambiguation method that finds the overlapping words in the dictionary definitions of the senses of two words. Word senses that have more overlapping words are marked as more similar. Satanjeev Banerjee [8] improved the Lesk Algorithm[7] and proposed an Adapted Lesk Algorithm[8].

The Adapted Lesk Algorithm[8] uses WordNet lexical database as a data source. While finding the similarity between two words, the algorithm searches the overlapping words or phrases in the word senses and also in the senses of the hypernym, holonym and hyponym relations of the words which are nouns and verbs. Only the attribute relation is used for the adjectives. Attribute relationship is declared as “*The attribute relationship links together noun and adjective synsets if the adjective synset is a value of the noun synset.*”[8]. For example, “*speed*” (noun) is the attribute of the “*fast*”(adjective) and also “*fast*”(adjective) is the attribute of the “*speed*” (noun), because of the symmetric relation.

The algorithm searches for the overlapping words or phrases in the glosses mentioned above. After counting the overlapped words or phrases, the relatedness is calculated by summing the square of each overlapping count.

Overlapping word count is used to calculate the similarity.

Adapted Lesk algorithm is too slow, because it looks to all words or phrases in the defined relations. The average computing time is shown in Table 1.

Adapted Lesk Tanimoto, David Hope[3] implemented a different similarity measure to the Adapted Lesk Algorithm[8]. The steps for calculating similarity between two words are:

- Creating the glosses of the words. Each gloss consists of hypernym, hyponym and holonym of the word. The algorithm also takes all the hyponym relations of the hypernyms of the word. In this gloss, words and the occurrence count of the words are stored.
- Generating a base gloss by combining the word glosses.
- Creating vectors for each gloss by comparing the word gloss and base gloss. If a word in the word gloss exists in the base gloss, then the occurrence count of the word is written to the vector. If not exists, then 0.0 is written to the vector.
- Applying the Tanimoto Similarity Distance[9] to the vectors. Tanimoto Similarity Metric is shown in Equation(1). v_1 and v_2 represents the vectors created by counting the occurrences of each word in the gloss.

$$T(w_1, w_2) = \frac{\sum v_1 * v_2}{v_1^2 + v_2^2 - \sum v_1 * v_2} \quad (1)$$

Adapted Lesk Tanimoto No Hyponyms, this algorithm applies the same method with Adapted Lesk Tanimoto Algorithm explained above. The only difference between these algorithms is that this algorithm does not take the hyponym relations of the hypernyms of the words.

Hirst-St.Onge, Hirst and St-Onge[10] introduced a similarity measurement algorithm that is implemented using the lexical chains proposed by Morris and Hirst[11]. David St-Onge [25] defined lexical chains as, “*A lexical chain is a succession of semantically related words in a text, that creates a context and contributes to the continuity of meaning.*”[12]

The algorithm calculates the relatedness of the words by looking the relations of the words in WordNet. A lexical chain is created by taking the gloss of the words and their relations in WordNet.

They have defined three different relations between words:

- Extra-strong relation points the words that are the same words,
- Strong relation shows the relation between compound names, words having the same synset and the words having the horizontal relation,
- Medium-strong relation represents a special path between words.

The weight of the words having extra-strong and strong relations is calculated with Equation (2) this formula C is a constant variable having the value 8.0.

$$weight = 2 * C \quad (2)$$

Medium-strong relations the weight is calculated with this formula, C and k are the constant variables.($C=8.0$ and $k=1.0$)

$$weight = C - pathlength - k * changes in directions \quad (3)$$

Resnik, Philip Resnik[13] proposed a method to find the semantic similarity using information content in a relational database. Similarity between two words is measured with this formula:

$$Sim(w1, w2) = \max_{c \in LCS(c1, c2)} [-\log P(c)] \quad (4)$$

LCS is the Least Common Subsumer that subsumes both word1 and word2. P represents the probability of concepts. Probability is calculated with:

$$P(word) = \frac{frequency\ of\ the\ word}{count\ of\ the\ noun\ or\ verb\ root\ counts} \quad (5)$$

Information Content (IC) value is equal to the probability of the word. It is calculated using Equation (4-5). Noun and verb root counts are the constant values that give the total noun and verb

roots in WordNet. The noun root count value is 128767 and the root count value for verbs is 95935 for WordNet 3.0.

Jiang-Conrath, J. Jiang and D.Conrath [14] proposed a similarity measurement method which is a combination of node-based approach[15] and edge-based approach[16] to get better results.

The approach uses both WordNet and WordNet Info-Content Files. WordNet Info-Content Files gives the information content derived from various corpus files like British National Corpus (BNC), Semcor and etc. This algorithm uses Semcor files and works only for nouns and verbs. Firstly, the Information Content(IC) value of the word is calculated to find the probability of the word in the WordNet Info-Content Files. The probability function is the fraction of word frequency to total verb/noun words count as shown in Equation (6).

$$P(word) = \frac{\text{frequency of the word}}{\text{count of the POS root counts}} \quad (6)$$

IC is calculated with the equation shown in Equation (7).

$$IC = -\log P(word) \quad (7)$$

While comparing two words, IC value is calculated for both words. Then the LCS is calculated by looking to all hypernyms of two words. IC values of the hypernyms are calculated and the one having the highest value is signed as LCS. The similarity function is :

$$Sim(w_1, w_2) = \frac{1}{IC(w_1) + IC(w_2) - 2 * IC(LCS(c_1, c_2))} \quad (8)$$

If two words are the same, the result would be infinity with Equation (8). Then the similarity is calculated with a different equation that is shown in Equation (9). “rootsum” represents the sum of the total count of the roots for nouns or verbs.

$$Sim(w_1, w_2) = \frac{1.0}{-\log((\text{rootsum} - 0.01) / \text{rootsum})} \quad (9)$$

Leacock-Chodorow, Leacock and Chodorow [17] proposed a similarity measure that is shown in Equation (10).

$$Sim(w_1, w_2) = -\log \frac{\text{length}(w_1, w_2)}{2 * D} \quad (10)$$

In the formula, length represents the shortest path between two words. For instance, length of the two siblings having the same root is 3. D represents the maximum depth in the taxonomy. Maximum depth for WordNet 3.0 is 19 for nouns and 12 for verbs.

Lin, Dekang Lin [18] defined a similarity measure. Lin approach has three definition for word similarity:

- As two words have more common they are more similar,
- Difference between two words indicates the dissimilarity between these words,
- Maximum similarity value is obtained if two words are the same,

The similarity is computed with this formula:

$$Sim(w_1, w_2) = \frac{2 * \log P(LCS(w_1, w_2))}{\log P(w_1) + \log P(w_2)} \quad (11)$$

Path, This measure calculates the similarity by getting the inverse of the length between two words. Length is the shortest path between two synsets.

Wu-Palmer, Wu and Palmer[19] defined a similarity measure defined as:

$$Sim(w_1, w_2) = \frac{2 * N_1}{N_2 + N_3} \quad (12)$$

In this definition, N_3 is the path from LCS to the root. N_1 and N_2 represent the path to the root for the $word_1$ and $word_2$ respectively. This measurement counts the nodes not edges between elements. The depth of a word is calculated by counting nodes from word to the root.

3. EVALUATION

This chapter introduces experimental data used for evaluation of word sense similarity algorithms and evaluation results. We have two different data sets for word sense disambiguation. The results will be given for both of them.

For each algorithm, computing time is calculated by taking the average of the hundred times execution of the similarity value between two words. These time values are shown in Table 1.

Table 1. *Computing time of the similarity algorithms*

| Algorithm | For two words(mSec) |
|-----------------------------------|----------------------------|
| Adapted Lesk | 72752,95 |
| Adapted Lesk Tanimoto | 5400,48 |
| Adapted Lesk Tanimoto No Hyponyms | 1865,92 |
| Hirst-StOnge | 3301,32 |
| Jiang-Conrath | 0,81 |
| Leacock-Chodorow | 1,7 |
| Lin | 3,09 |
| Path | 2,17 |
| Resnik | 0,82 |
| Wu-Palmer | 4,29 |
| Disco | 6,19 |

As shown in Table 1, some of “the algorithms” have long computing time and not included in the evaluation process. These algorithms are Adapted Lesk, Adapted Lesk Tanimoto, Adapted Lesk Tanimoto No Hyponyms, and Hirst-StOnge algorithms.

3.1 *Experimental Data*

For evaluation two different datasets has been used. These datasets are taken from Keymantic[1] work and named Unimore and IMDB. Unimore is a schema related with a university. IMDB is related with cinema. These datasets consists of keywords and xml schemas. For calculating similarity between words, we have collected words from keywords and xml schema elements. Keywords represent the possible words or phrases that people query in an information retrieval system. We extracted words from these keywords and compared them with schema elements and attributes. Aim of this comparison is finding the similarity between keywords and schema elements for disambiguating the words.

While constructing word collection, we make the plural words singular, because the algorithms working with WordNet can not find the plural words in the thesaurus. In addition, abbreviations like “prof.” and “prof” are deleted from the collection. We obtained 44 keywords and 38 attributes from Unimore database, 20 keywords and 54 attributes from IMDB database.

3.2 *Intervals*

For the similarity methods evaluation, we calculate the similarity value of the keywords and the data elements taken from the Unimore and IMDB datasets. These values are written in a matrix in which each cell points the similarity value of the corresponding keyword and attribute.

Three intervals are defined for evaluation. The reason of the assigning the intervals is to use only the values that shows the maximum similarity. Firstly, maximum similarity value is calculated for each metric. By multiplying this maximum value with 0.9, 0.8 and 0.7, three intervals are obtained.

All of the similarity metrics in JWS give the same value for the identical word. Only Resnik [13] method gives different values for identical words. Because this method uses LCS value to calculate the similarity. Therefore, similarity value for the same words is different from the other words.

3.3 Similarity Metrics

For evaluating similarity values, two similarity metric is defined. Similarity 1 shows the similarity between the same words or synonyms. It shows an IS-A relation, like the relation between “movie” and “film”. Similarity 2 shows hyponym or hypernym relations between the words. For evaluating the algorithms, we defined these metrics and created a template that shows the word pairs that are expected to be similar.

3.3 Experimental Results

A template for evaluation is created based on human similarity judgements. The template is a matrix that shows similar words by giving “1” value for corresponding keyword and data element. Otherwise the value is empty. The results calculated by matching the template and the similarity matrix for each interval. The values in the intervals are counted as true results. For calculating the precision of the algorithms, these values are compared with the expected results. Precision values are separately calculated for Similarity1 and Similarity2 in two different datasets Unimore and Imdb. Precision value is the ratio of the number of true matches in that interval to the number of the total values in that interval. In Table 2, Table 3, Table 4 and

Table 5, the precision value for each algorithm with Sim1 and Sim2 is given with different datasets.

Table 2. Precision value of Sim1 for Unimore dataset

| Similarity Metric Sim1 – Unimore | Interval | | |
|-------------------------------------|----------|-------|-------|
| | 1 | 2 | 3 |
| Jiang-Conrath | 89,47 | 0 | 0 |
| Leacock&Chodorow | 100 | 21,42 | 2,77 |
| Wu&Palmer | 56,09 | 1,36 | 1,67 |
| Path | 100 | 0 | 0 |
| Lin | 85,71 | 0 | 10 |
| Resnik | 60 | 37,5 | 57,14 |
| DISCO | 100 | 0 | 0 |

Table 3. Precision value of Sim2 for Unimore dataset

| Similarity Metric Sim2 - Unimore | Interval | | |
|-------------------------------------|----------|------|----|
| | 1 | 2 | 3 |
| Jiang-Conrath | 100 | 0 | 0 |
| Leacock&Chodorow | 90 | 0 | 0 |
| Wu&Palmer | 40 | 0 | 0 |
| Path | 100 | 0 | 0 |
| Lin | 9,52 | 20 | 20 |
| Resnik | 0 | 12,5 | 0 |
| DISCO | 0 | 0 | 0 |

Table 4. Precision value of Sim1 for IMDB dataset

| Similarity Metric Sim1 - IMDB | Interval | | |
|----------------------------------|----------|-------|-----|
| | 1 | 2 | 3 |
| Jiang-Conrath | 0 | 0 | 0 |
| Leacock&Chodorow | 0 | 0 | 0 |
| Wu&Palmer | 4 | 25,5 | 1,8 |
| Path | 100 | 0 | 0 |
| Lin | 66,6 | 0 | 10 |
| Resnik | 33,3 | 66,66 | 20 |
| DISCO | 100 | 0 | 0 |

Table 5. Precision value of Sim2 for IMDB dataset

| Similarity Metric Sim 2 – IMDB | Interval | | |
|-----------------------------------|----------|-------|------|
| | 1 | 2 | 3 |
| Jiang-Conrath | 89,47 | 0 | 0 |
| Leacock&Chodorow | 100 | 21,42 | 2,77 |
| Wu&Palmer | 56,09 | 1,36 | 1,67 |
| Path | 100 | 0 | 0 |
| Lin | 0 | 0 | 0 |
| Resnik | 0 | 0 | 10 |
| DISCO | 0 | 0 | 0 |

All the results are listed in the tables for each algorithm with different intervals. Figure 1 illustrates the Sim1 results of all algorithms with Unimore dataset in the first interval. Leacock-Chodorow (L&C), Path and DISCO have the maximum precision with 100% value. Jiang-Conrath (J&C) and Lin methods have precision value more than 85%. Resnik and Wu-Palmer (Wup) have lower performance than the others with a precision value between 56% and 60%.

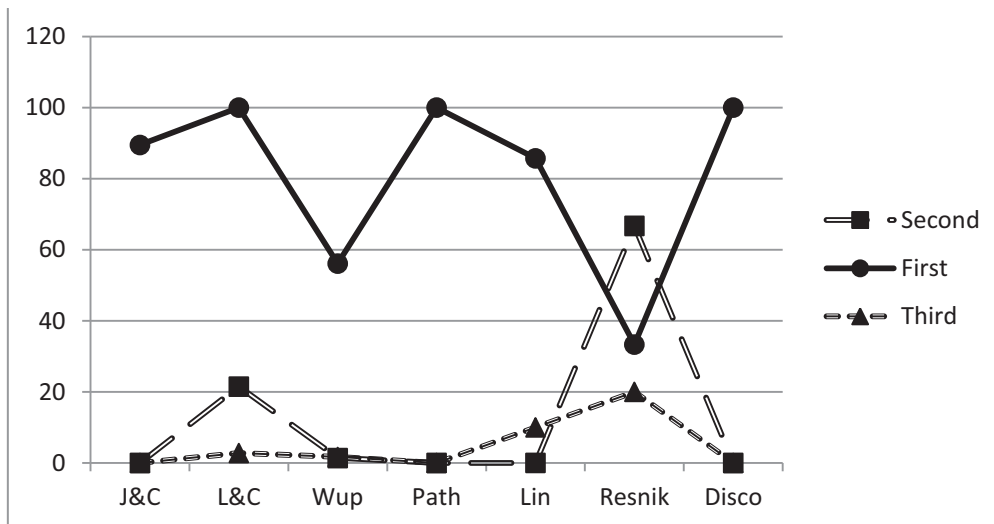


Figure 1. Results for Sim1 with all algorithms using Unimore dataset

For the second and third interval, most of the algorithms give values around zero. In these intervals, L&C and Resnik algorithms that use IC give higher results. With this result, we may have higher results with the lower similarity values by using probability of the word senses (IC).

The results for Sim2 with Unimore dataset is shown in Figure 2. The results have lower precision value than Sim2. L&C metric has the maximum precision value for Sim2 in the second and third interval. Lin method has 20% precision in the second and third interval. The algorithms find similarity for Sim2 with lower values.

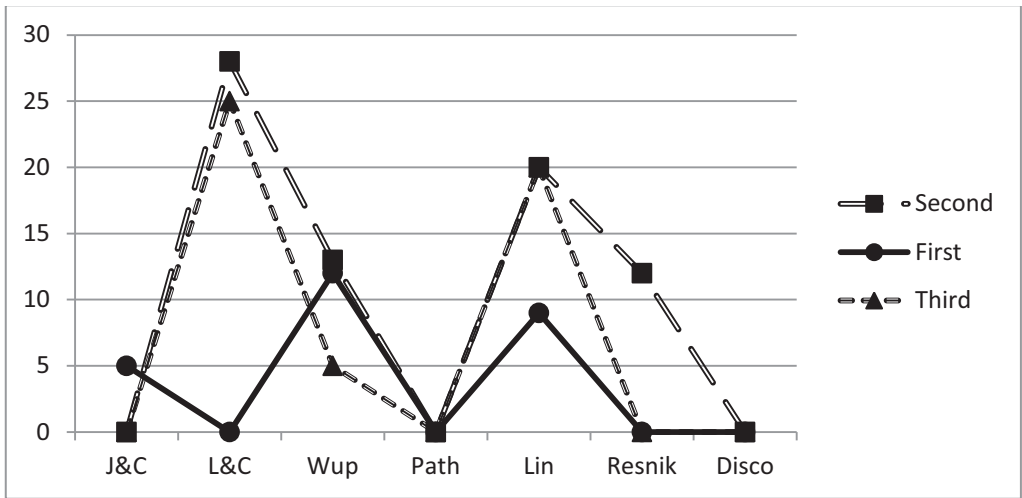


Figure 2 Results for Sim2 with all algorithms using Unimore dataset

Sim1 results with Imdb dataset are shown in Figure 3. The methods having the maximum precision value (100%) are the J&C and Path. L&C also has a high precision value with 90%. Sim1 results with this dataset is different from the Unimore dataset. Wup, Lin and Resnik metrics give lower precision values then the values with Unimore dataset.

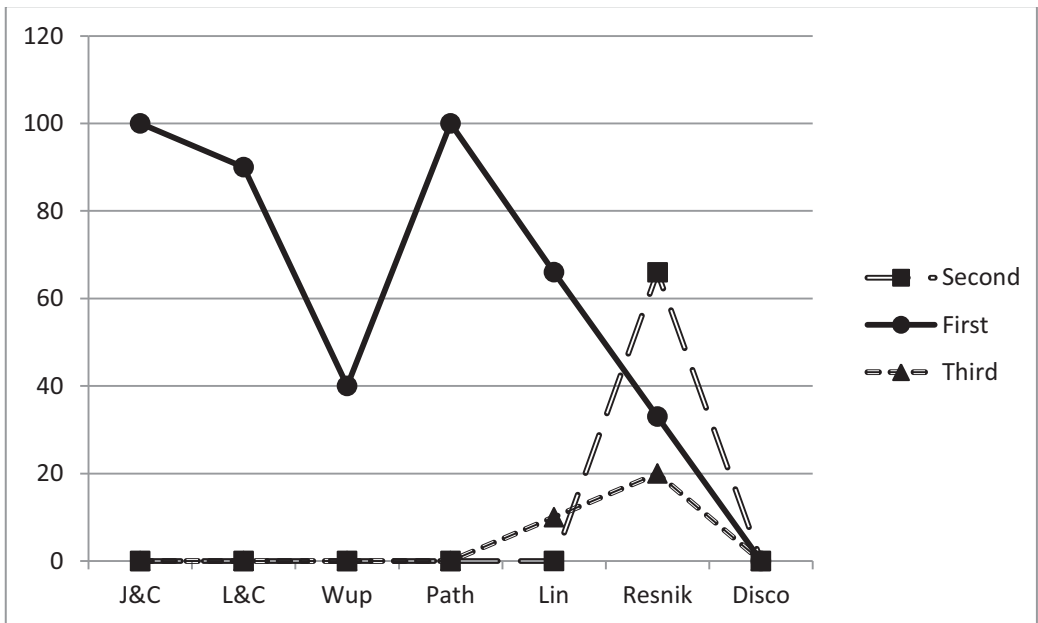


Figure 3 Results for Sim1 with all algorithms using Imdb dataset

Figure 4 illustrates the results for Sim2 with Imdb database. As seen in the chart most of the algorithms have no similar values in all intervals. Only Wup gives results in all intervals. Resnik gives 10% precision value in third interval.

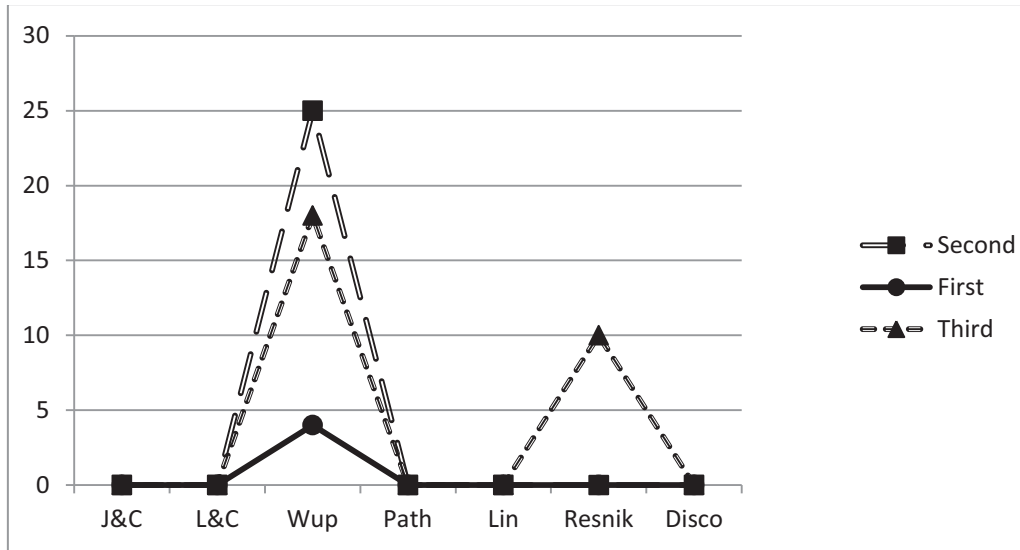


Figure 4 Results for Sim2 with all algorithms using Imdb dataset

4. CONCLUSIONS

We can see that the algorithms using the path between the word senses are more successful than the others –considering the algorithms in JWS. Wup uses LCS beside the path between the word senses and has lower performance than the others considering only the path. If we look the algorithms using the IC for measuring similarity, we can easily see that they have minimum 60% precision value with Resnik method. Algorithms that use not only IC but also LCS have higher precision values (J&C with 89% and Lin with 85%).

As a conclusion, J&C, L&C, Path and DISCO similarity measures have higher precision values (minimum 89%) for Sim1 with both datasets. For Sim2, J&C, L&C and Path give higher precision with Unimore dataset.

5. REFERENCES

- [1] Sonia Bergamaschi, Elton Domnori, Francesco Guerra, Mirko Orsini, Raquel Trillo Lado, Yannis Velegrakis, "Keymantic: Semantic Keyword-based Searching in Data Integration Systems," *PVLDB*, vol. 3, no. 2, pp. 1637-1640, 2010.
- [2] "Merriam-Webster Dictionary," [Online]. Available: <http://www.merriam-webster.com/>. [Accessed 08 2014].
- [3] D. Hope, "David Hope, Sussex University, Cognitive and Language Processing Systems Group (CALPS)," University of Sussex, Brighton, England, BN2, [Online]. Available: <http://www.sussex.ac.uk/Users/drh21/>. [Accessed 19 March 2013].
- [4] P. Kolb, " DISCO: A Multilingual Database of Distributionally Similar Words," in *KONVENS-2008*, Berlin, 2008.
- [5] P. Kolb, "Experiments on the difference between semantic similarity and relatedness," in *17th Nordic Conference on Computational Linguistics - NODALIDA '09*, Odense,Denmark, 2009.
- [6] T. Pedersen, "Ted Pedersen - WordNet::Similarity," [Online]. Available: <http://wn-similarity.sourceforge.net/>. [Accessed 19 March 2013].
- [7] M. Lesk, "Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone," in *SIGDOG'86*, 1986.
- [8] S. Banerjee, «Adapting the Lesk Algorithm for Word Sense Disambiguation to WordNet,» University of Minnesota, Duluth, Minnesota, 2002.
- [9] "Jaccard index - Wikipedia, the free encyclopedia," Wikimedia Foundation, Inc., [Online]. Available: http://en.wikipedia.org/wiki/Jaccard_index#Tanimoto_Similarity_and_Distance. [Accessed 9 5 2013].
- [10] Graeme Hirst, David St-Onge, "Lexical chains as representations of context for the detection and correction of malapropisms," Fellbaum, 1998.
- [11] "Lexical cohesion computed by thesaural relations as an indicator of the structure of text,"

Computational Linguistics, vol. 17, pp. 21-48, 1991.

- [12] D. St-Onge, "Detecting and Correcting Malapropisms with Lexical Chains," Toronto, 1995.
- [13] P. Resnik, "Using Information Content to Evaluate Semantic Similarity in a Taxonomy," in *14th International Joint Conference on Artificial Intelligence*, Montreal, 1995.
- [14] Jay J. Jiang, David W. Conrath, in *International Conference Research on Computational Linguistics (ROCLING X)*, Taiwan, 2007.
- [15] P. Resnik, "WordNet and Distributional Analysis: A Class-based Approach to Lexical," in *AAAI Symposium on Probabilistic Approaches to Natural*, San Jose, CA, 1992.
- [16] Rada, R., H. Mili, E. Bicknell, and M. Bletner, "Development and Application of a Metric," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 1, pp. 17-30, 1989.
- [17] Claudia Leacock, Martin Chodorow, "Combining local context and wordnet similarity for word sense identification," *Fellbaum*, pp. 265-283, 1998.
- [18] D. Lin, "An information-theoretic definition of similarity," in *15th International Conference on Machine Learning*, San Francisco, CA, 1998.
- [19] Zhibiao Wu, Martha Palmer, "Verb semantics and lexical selection," in *32nd Annual Meeting of the Association for Computational Linguistics*, 1994.

Energy Optimization of the Current-Limiting Power LED Drivers in Various Powers

Ö. F. Farsakoğlu¹, H. Y. Hasırcı², İ. Atik³, İ. Çelik⁴

^{1,2,3,4}Electrical and Electronics Engineering Department, Kilis7 Aralık University, Faculty of Engineering and Architecture, 79000, Kilis, TURKEY

¹ffarsakoglu@kilis.edu.tr, ²hyusufhasirci@kilis.edu.tr, ³ipekinal@kilis.edu.tr,
⁴ibrahimcelik@kilis.edu.tr

ABSTRACT

As the world population increases, the energy consumption also increases. The sources of providing energy needs are decreasing step by step. Besides the necessity of finding alternative energy sources, it is also important to make savings and use more efficient systems. So, the lighting element used should be selected more efficiently. At the present, power LEDs sizes are smaller, more durable, and more long life than traditional light sources. Besides, power LEDs also have more electrical, thermal, and optical efficiencies than other lighting systems. All these circumstances are strongly preferable for designers and investors. Power LEDs have high current values. For this reason, special drivers are needed. Power LED drivers are classified as voltage limiting and current limiting drivers. In this study, current limiting LED drivers were investigated at low power conditions. In this case, driver circuit designs have been carried out in terms of the LED luminous efficiency optimization. Driver circuits were designed at the power of 1W, 3W, and 5W. Then, the data of energy efficiencies of these circuits were tabulated as a comparison.

Keywords: lighting systems, power LEDs, current limiting LEDs drivers, energy efficiency, optical efficiency.

1. GİRİŞ

Türkiye’de birincil enerji tüketimindeki artış oranı yıllık % 4,3 olarak açıklanmış olup, enerji ihtiyacının yüzde 75’lik bölümü dış kaynaklardan karşılanmaktadır. Yüzde 90’dan daha fazla gibi bir değerde ithal edilen doğal gazın yaklaşık yarısı elektrik enerjisine çevrilerek kullanılmaktadır. Gelişmiş ve gelişmekte olan ülkelerde genel enerji tüketimindeki artış, elektrik enerjisi tüketimini de arttırmaktadır. Bu doğrultuda artan tüketim talebinin karşılanması için ülkemizdeki mevcut kurulu gücün önümüzdeki 5 ile 10 yıl arasında en az iki katına çıkması gerektiği öngörülmektedir.

Aydınlatma sistemlerinde harcanan enerji, ülkemizde tüketilen elektrik enerjisinin yaklaşık % 20’sine karşılık gelmektedir. Bu oran kullanım alanlarına göre konutlarda % 20, sanayide % 10, ofis gibi ticari yapılarda ise en az % 30 olarak değişmektedir.

Aydınlatmada gerçekleştirilebilecek enerji tasarrufunun önemi, özellikle elektrik enerjisi üretiminde yaşanan sorunlar göz önüne alındığında kendiliğinden ortaya çıkmaktadır. Tüm enerji verimlilik çalışmalarında olduğu gibi aydınlatmada da tasarrufun, çalışma emniyeti ve konfor koşullarından ödün verilmeden yapılması gerekmektedir. Aydınlatma sistemleri her şeyden önce ortamda yapılacak işe uygun görme koşullarının konforlu bir şekilde sağlanabilmesi amacı ile tesis edilmelidir. Daha önceleri aydınlatma tesisatlarında doğru eleman seçimi ve doğru tasarımların yapılmamasının yanı sıra gelişen teknoloji ile son yıllarda yüksek verimliliklere sahip LED’lerin

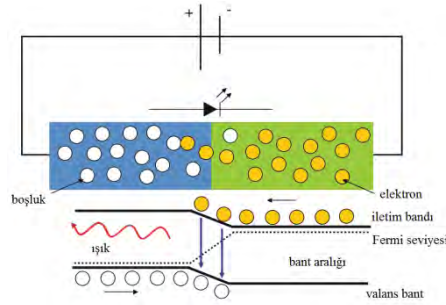
(LightEmittingDiode -Işık Yayan Diyot) ortaya çıkması, şu anda hem enerji tasarruf potansiyel oranı yüksek, hem de geri ödeme süreleri kısa olan enerji verimlilik çalışmaları olarak öncelikli bir şekilde karşımıza çıkmaktadır[1].

Standart 5 mm LED'ler 20mA gibi çok küçük akımlarla çalışırlar[2]. Fakat LED'lerin, konvansiyonel aydınlatma elemanlarından olan akkor lambalar, floresan lambalar, halojen lambalar gibi ışık kaynaklarına iyi birer alternatif olmaları için daha yüksek ışık veren çeşitleri üretilmiştir. Bu LED'ler, güç LED'i(power LED) olarak adlandırılıp, 350mA, 700mA, 1050mA ve hatta daha yüksek akım çekmektedirler[3]. Bu sebepten dolayı sabit akım çeken özel sürücülere ihtiyaç duyulmaktadır.

Bu çalışmada, düşük güçlerde akım sınırlamalı LED sürücülere üzerinde inceleme yapılmıştır. LED ışık verimliliğinin optimizasyonu yönünden düşük güçteki sürücü devrelerinin tasarımları gerçekleştirilmiştir. Müteakiben bu devrelerin verimlilikleri belirlenerek karşılaştırma tabloları oluşturulmuştur.

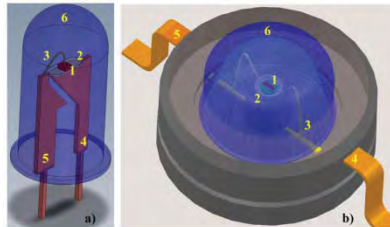
2. LED'in (LIGHT EMITTING DIODE - IŞIK YAYAN DİYOT) YAPISI

Elektrik enerjisini ışık enerjisine dönüştürmek için tasarlanmış yarı iletken elemanlar LED (ışık yayan diyotlar) diye adlandırılırlar. LED'in yapısı ve elektriksel davranışları bir doğrultucu diyota benzemektedir. LED'ler doğrultucu diyotlar gibi yarı iletken p-tipi ve n-tipi malzemenin birleşimi ile yapılmıştır. Bu malzeme anot ve katot olarak adlandırılan iki terminalle bağlanır. LED'in anodu katoduna göre pozitif olduğunda üzerinde elektrik akımı geçecektir. Bu durumda LED elektrik enerjisini ışık enerjisine dönüştürerek ışık yaymaya başlayacaktır. Fakat tersi durumda LED çalışmamaktadır [4]. Şekil 1'de LED'in çalışma prensibi gösterilmiştir.



Şekil 1. LED'in Çalışma Prensibi[5]

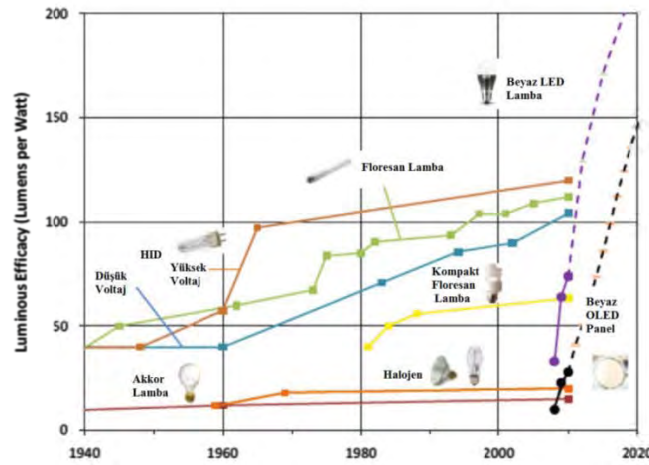
Güç LED'leri yüksek lümen değerine sahip olması nedeniyle aydınlatma sektöründe yaygın olarak kullanılmaktadır. Güç LED'lerinin akım seviyeleri standart 5 mm çaplı LED'lerden çok daha yüksektir. Şekil 2'de standart 5 mm LED ve güç LED'inin parçaları gösterilmiştir.



Şekil 2.a) Standart 5 mm LED b) Güç LED basitleştirilmiş bir gösterimini yansıtır[6]

1: aktif çip, 2: genellikle küçük reflektör, katot kurşun yerleştirilir, 3:altın tel, 4: anot uçları,5: katot uçları 6: aktif çipin korunması ve yayılan ışığı toplamak için saydam bir malzemeden yapılmış lens şeklinde bir kapsül [6].

LED'lerin diğer aydınlatma elemanlarına göre daha çok tercih edilmelerinin sebepleri,LED aydınlatma sistemlerinin daha yüksek enerji verimliliği sağlaması,LED'in uzun ömürlü olmasıyaklaşık olarak 50 000 ile 100000 saat arasında ve bu nedenle daha az bakım gerektirmesi, ekipmanların boyutunun küçük olması bu nedenle az yer kaplamaları, ışık ve renk çeşitliliği sağlamaları, darbeler ve titreşimlere daha dayanıklı olmaları ve daha düşük güç tüketimleri gibi buna benzer birçok avantajlara sahiptir [7].Bu avantajlardan dolayı LED ışık kaynaklarının önümüzdeki 10 yılda konvansiyonel aydınlatma elemanları ile yer değiştirmesi ön görülmektedir. Şekil 3'te çeşitli ışık kaynaklarının aydınlanma verimlilikleri gösterilmiştir.



Şekil 3. Işık kaynaklarının aydınlanma verimliliği [8].

Şekil 3'te de görüldüğü gibi akkor lambanın ışık verimliliği 15-20 lm/ W [9], floresan lambanın 50-100 lm/W, kompakt floresan lambanın 50-70 lm/W iken LED'lerin ışık verimliliği 80-160 olduğu hatta 2020 yılında bu verimliliğin daha da artacağı görülmektedir [8]. Tablo 1 de bazı ışık kaynaklarının verimliliği ve ömürleri gösterilmektedir.

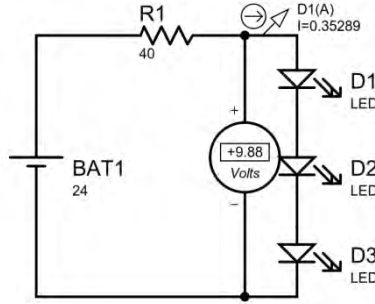
Tablo 1. Bazı ışık kaynaklarının aydınlanma verimliliği ve ömürleri [8].

| Işık Kaynağı | Aydınlanma Verimi (lm/W) | Lamba Ömrü (saat) |
|------------------------|--------------------------|-------------------|
| Akkor | 15-20 | 1000 |
| Tungsten Halojen | 12-35 | 2000-4000 |
| Cıva Buharlı | 40-60 | 12000 |
| Kompakt Floresan Lamba | 40-70 | 6000-12000 |
| Floresan Lamba | 50-100 | 10000-16000 |
| İndüksiyon lambası | 60-80 | 60000-100000 |
| Metal Halojen | 50-100 | 6000-12000 |
| Yüksek Basınçlı Sodyum | 80-100 | 12000-16000 |
| LED | 80-160 | 50000-100000 |

3. AKIM SINIRLAMALI LED SÜRÜCÜLERİ

3.1 Bir Dirençle Yapılan Basit Sürücü Devresi

LED'ler ışık yaymak için sabit akıma ihtiyaç duyarlar. Şekil 4'te sadece bir direnç kullanılarak yapılan basit sürücü devresi gösterilmiştir.



Şekil 4. Basit sürücü devresi

Şekil 4'deki devrede 24V'luk sabit bir gerilim kaynağı, 3 tane 1 W'lık güç LED'leri ve akımı sınırlamak için de bir direnç bulunuyor. Devreye ilk enerji verildiğinde LED'ten 0,35A akım geçer. Zaman geçtikçe LED'ler ısınmaya başlar ve LED'lerin ileri yön gerilimi düşer. Bu durumda Ohm - Yasası gereğince devreden geçen akım artar ve sıcaklık belirli bir dereceye ulaştığında LED'e uygulanan akım 0,39 A düzeyine çıkar. Akımın artışı LED'in daha çok ısınmasına neden olur ve önlem alınmazsa, LED bir süre sonra bozulabilir.

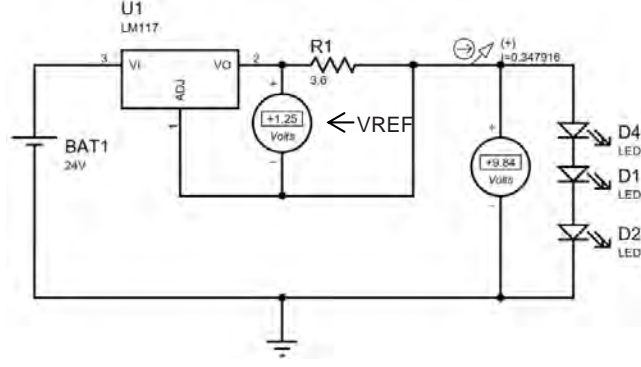
Basit bir sürücü devrede sürücü akımı,

$$I = \frac{V_{kaynak} - V_f * n}{R} \quad (1)$$

bağıntısı ile hesaplanır. Burada I, LED'e uygulanan akımı, V_{kaynak} , devreye uygulanan gerilim değerini, V_f , LED'in ileri yön gerilimini, n ise LED sayısını göstermektedir.

3.2 LM117 Regülatör Devresi ile Yapılan LED Sürücüsü

Diğer bir LED sürme devresi de LM117 entegre devresi ile yapılan sürücü devredir. Giriş gerilimi 1.2V ile 30 V arasında ve çıkış akımı 1.5 A' e kadar kullanılabilir [10]. LM117 yüksek güçteki LED sürücüleri için kullanılmaya uygun değildir. Şekil 5'teki gibi adjust ucuna bağlanan dirençler sayesinde çıkış gerilimini sabit tutar. Bu sürücüde kullanılan LM117'te devre iletiminde yaklaşık olarak 4.5 W kayıp olmaktadır. Bu da, anılan sürücü devresinin verimini düşürmektedir.



Şekil 5. LM117 ile yapılan sürücü devresi.

Şekil 5’deki devrede çıkış akımı I_{out} ve Referans gerilimi V_{REF} ,

$$I_{out} = \frac{V_{REF}}{R_1^*} \quad (2.a)$$

$$V_{REF} = 1.25 \text{ V}, \quad R_1^* = 0.8 \leq R1 \leq 120 \quad (2.b)$$

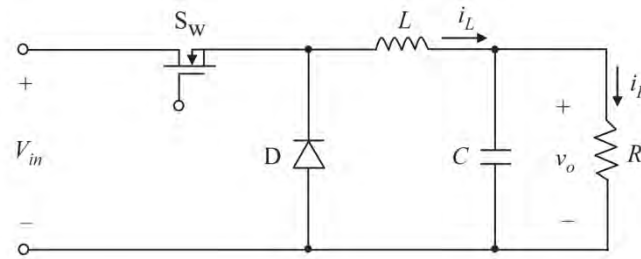
bağıntıları ile verilir.

3.3 DC-DC Düşürücü (Buck) Konvertör

Düşük güçteki akım sınırlamalı LED sürücülerinin genellikle düşürücü (buck) tipleri kullanılmaktadır. Bunun sebebi giriş gerilimleri çıkış gerilimlerinden yüksek olmasıdır. Şekil 6’da düşürücü konvertörlerin devre şeması verilmiştir. Şekil 6’da devre üzerinde V_{in} giriş gerilimini, S kontrol anahtarını, D diyodu, L filtreleme indüktörünü, C filtre kapasitörünü ve R de yükü temsil etmektedir [11]. Bu parametrelerle ilgili bağıntılar aşağıda verilmiştir.

$$\frac{di_L}{dt} = \frac{1}{L}(V_{in} - V_0) \quad (3)$$

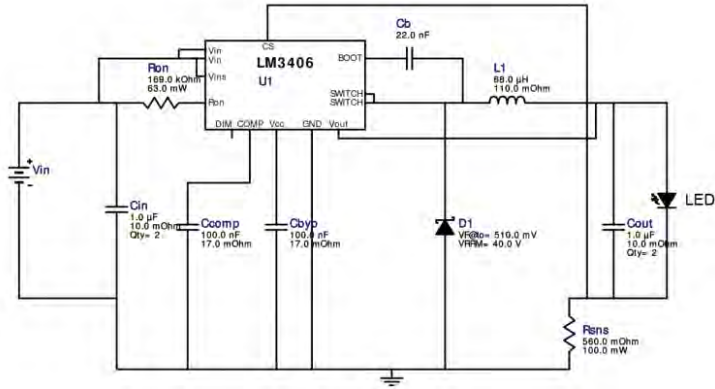
$$\frac{dV_0}{dt} = \frac{1}{C}(i_L - \frac{V_0}{R}) \quad (4)$$



Şekil 6. DC-DC düşürücü konvertör[11]

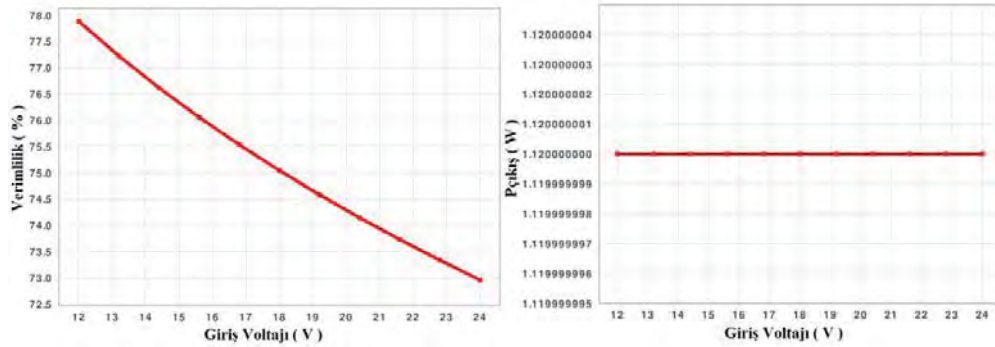
3.3.1 LM3406 ile Yapılan Akım Sınırlamalı Güç LED Sürücüsü

LM3406 gerilim düşürücü bir regülatördür. Bu, düşük voltaj referansı, geniş giriş voltaj aralığı ve hızlı bir çıkış fonksiyonuna sahiptir [12]. LED'ler için sabit bir akım sağlamanın yanı sıra 1.5A'ye kadar yüksek akım gerektiren yerlerde de kullanılmaktadır.

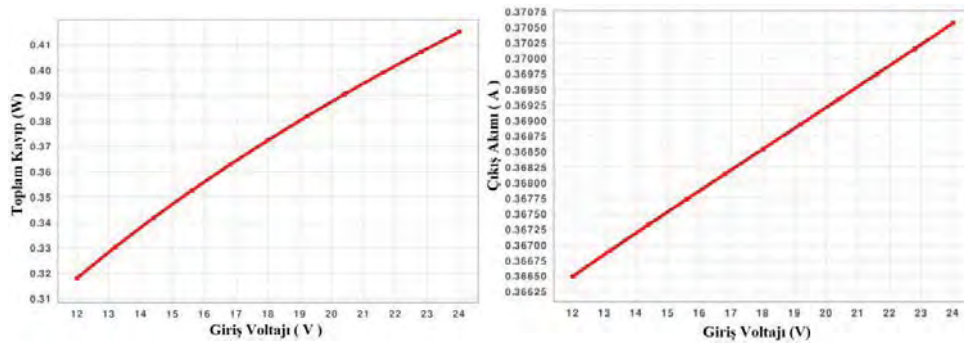


Şekil 7. 1 W'lık Güç LED sürücüsü

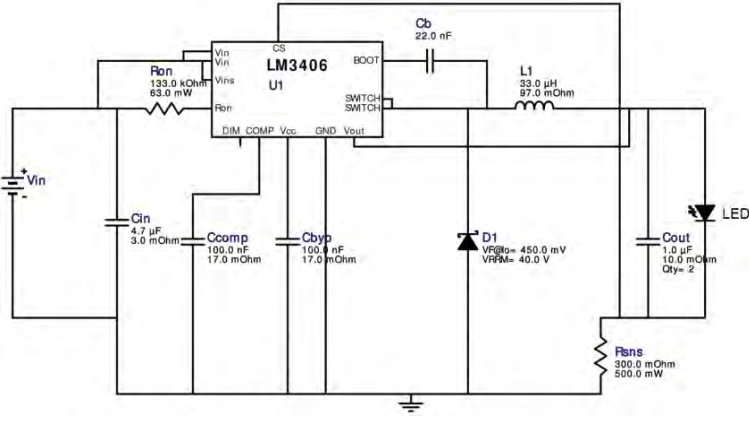
LM3406 ile yapılan sürücü devresinde giriş gerilimi 12 V ile 24 V arasında değişirken çıkış akımı 1 Watt LED sürücü için 350 mA (0.35 A). Şekil 7'deki devre 1 W'lık LED sürücü devresi gösterilmiştir. Aşağıda devrenin verimlilik, çıkış gücü (Pout), toplam kaybı ve LED'in çıkış akımlarının değişimleri verilmiştir.



Şekil 8. 1 W'lık Güç LED sürücüsünün verimliliği Şekil 9. 1 W'lık Güç LED sürücüsünün çıkış gücü

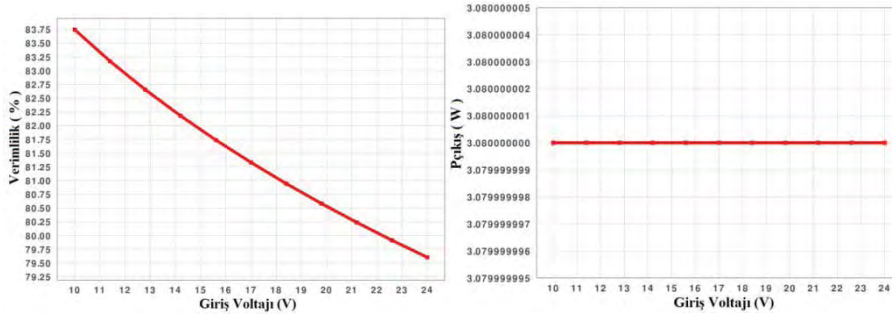


Şekil 10. 1 W'lık Güç LED sürücüsünde toplam kayıp Şekil 11. 1 W'lık Güç LED sürücüsünün çıkış akımı

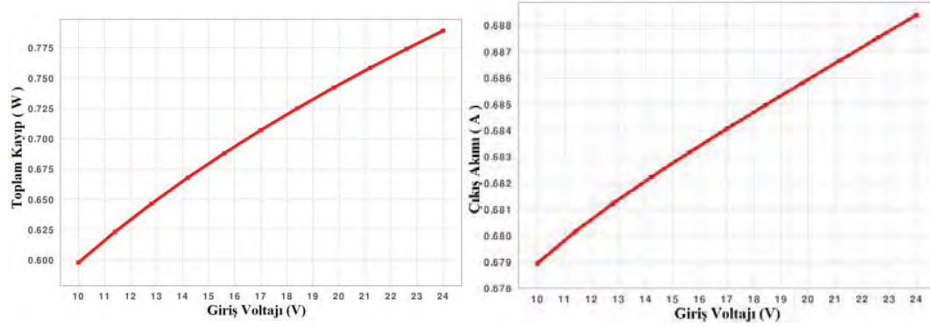


Şekil 12.3 W'lık Güç LED sürücüsü

LM3406 ile yapılan sürücü devresinde minimum giriş gerilimi 10 V ile maksimum giriş gerilimi 24 V dur. Bu değerler arasında çıkış akımı 3Watt LED sürücü için 700 mA (0.7 A)dir.Aşağıda devrenin verimlilik, çıkış gücü (Pout), toplam kaybı ve LED'in çıkış akımlarının grafikleri verilmiştir.

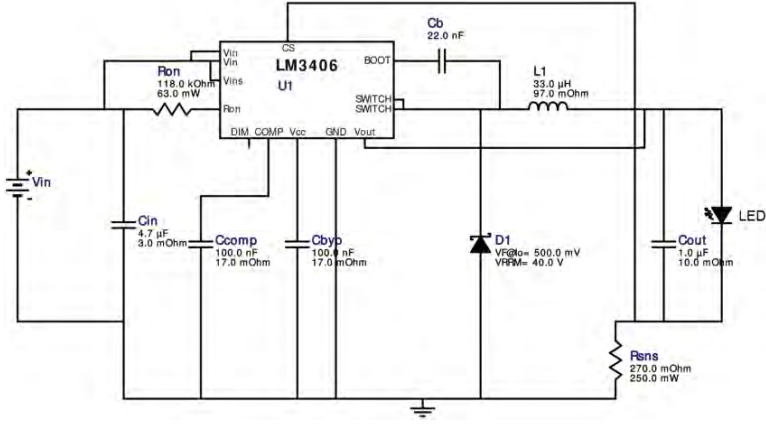


Şekil 13. 3 W'lık Güç LED sürücüsünün verimliliği Şekil 14.3 W'lık Güç LED sürücüsünün çıkış gücü



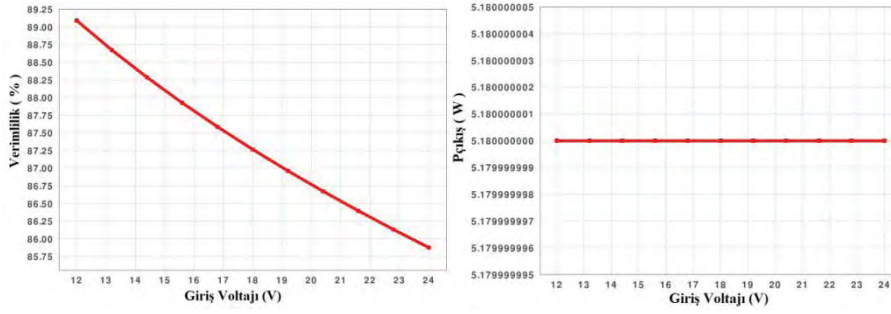
Şekil 15. 3 W'lık Güç LED sürücüsünde toplam kayıp

Şekil 16. 3 W'lık Güç LED sürücüsünün çıkış akımı

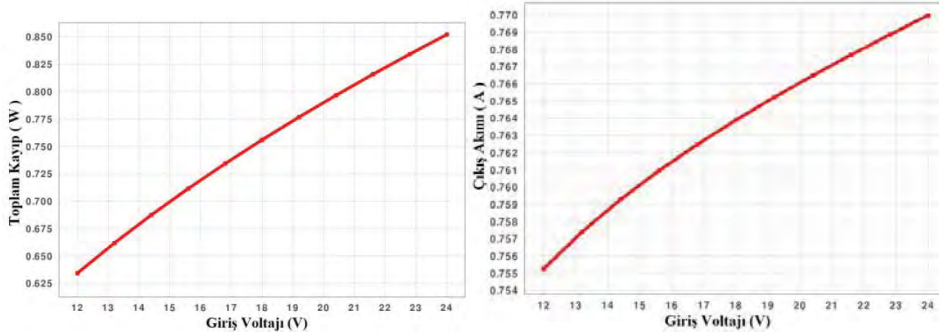


Şekil 17. 5 W'lık Güç LED sürücüsü

LM3406 ile yapılan sürücü devresinde minimum giriş gerilimi 12 V ile maksimum giriş gerilimi 24 V dur. Bu değerler arasında çıkış akımı 5Watt LED sürücü için 700 mA (0.7 A)dir. Aşağıda devrenin verimlilik, çıkış gücü (Pout), toplam kaybı ve LED'in çıkış akımlarının grafikleri verilmiştir.



Şekil 18. 5 W'lık Güç LED sürücüsünün verimliliği Şekil 19. 5 W'lık Güç LED sürücüsünün çıkış gücü



Şekil 20. 5W'lık Güç LED sürücüsünde toplam kayıp

Şekil 21. 5 W'lık Güç LED sürücüsünün çıkış akımı

Tablo2.Bir dirençle yapılan basit sürücü devresinin sonuçları

| LED Sürücünün Gücü | LED Sürücünün Giriş Gerilimi (V) | LED Sürücünün Verimliliği (%) | LED Sürücünün Çıkış Akımı(mA) | LED Sürücünün Çıkış Gücü (Watt) |
|--------------------|------------------------------------|---------------------------------|---------------------------------|---------------------------------|
| 1 Watt | 12 | 27 | 340 | 1.14 |
| | 24 | 13 | 342 | 1.12 |
| 3 Watt | 12 | 82 | 350 | 3.3 |
| | 24 | 41 | 350 | 3.4 |
| 5 Watt | 12 | - | - | - |
| | 24 | 68 | 340 | 5.6 |

Çıkış güçleri 1 W, 3 W ve 5W olan bu sürücü devreleri Proteus 8 programı kullanılarak tasarımları gerçekleştirilmiştir. Tasarımı yapılan bu devrelerin analizleri yapılmış ve sonuçlar Tablo2’de verilmiştir. Tablo 2’deki sonuçlar incelendiğinde giriş geriliminin düşük olduğunda verimin daha yüksek olduğu görülmektedir.

Tablo3. LM117 regülatör devresi ile yapılan LED sürücüsü devresinin sonuçları

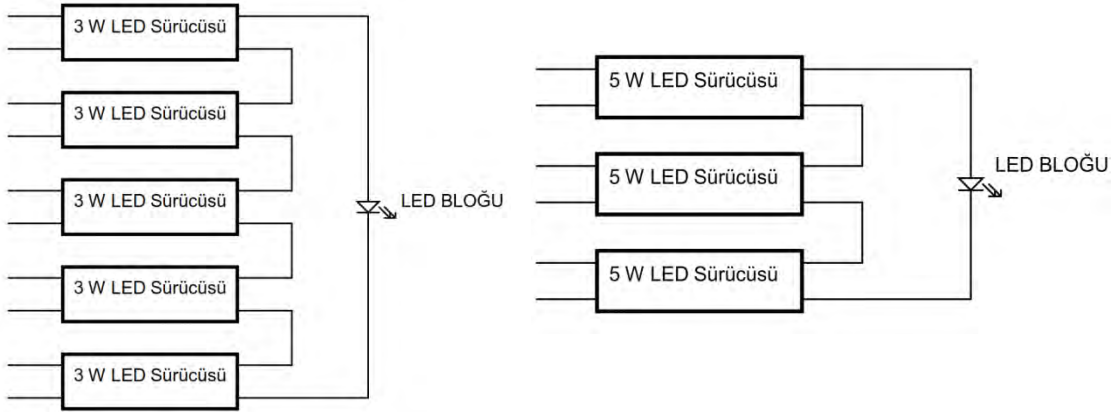
| LED Sürücünün Gücü | LED Sürücünün Giriş Gerilimi (V) | LED Sürücünün Verimliliği (%) | LED Sürücünün Çıkış Akımı(mA) | LED Sürücünün Çıkış Gücü (Watt) |
|--------------------|------------------------------------|---------------------------------|---------------------------------|---------------------------------|
| 1 Watt | 12 | 27 | 350 | 1.15 |
| | 24 | 14 | 350 | 1.15 |
| 3 Watt | 12 | 82 | 344 | 3.2 |
| | 24 | 41 | 345 | 3.3 |
| 5 Watt | 12 | - | - | - |
| | 24 | 68 | 355 | 5.4 |

LM117 regülatör devresi ile yapılan LED sürücüsü devresi farklı güçlerde aynı şekilde Proteus 8’de tasarlanmıştır. Tasarımı yapılan bu devrelerin analizleri yapılmış ve sonuçlar Tablo3’de verilmiştir. Verilerin incelenmesinden düşük gerilim değerlerinde verimliliğin daha yüksek olduğu görülmektedir. Kullanılan LM117 entegresiyaklaşık olarak 4 W- 5 W arası enerji harcadığı belirlenmiştir. Bunun ise verimi düşüren bir etken olduğu görülmektedir.

Tablo4. LM3406 ile yapılan sürücü devresinin sonuçları

| LED Sürücünün Gücü | Giriş Gerilimi (V) | Verimliliği (%) | Çıkış Akımı (mA) | Çıkış Gücü (Watt) | Toplam Kayıp (Watt) | LED Vf (ileri yön gerilimi(V) | Işık Akısı (lm) |
|--------------------|----------------------|-------------------|-------------------|-------------------|---------------------|---------------------------------|------------------|
| 1 Watt | 12 | 78.8 | 360 | 1.12 | 0.31 | 3.2 | 90-110 |
| | 24 | 73 | 370 | 1.12 | 0.42 | 3.2 | |
| 3 Watt | 12 | 83 | 680 | 3.08 | 0.63 | 4.2 | 160-190 |
| | 24 | 79.7 | 695 | 3.08 | 0.78 | 4.2 | |
| 5 Watt | 12 | 89.1 | 750 | 5.18 | 0.62 | 7.2 | 400 |
| | 24 | 86 | 770 | 5.18 | 0.85 | 7.2 | |

Düşürücükonverterle yapılan sürücü devresinin tasarımları WEBENCH Designer'da yapılmıştır. Yapılan analizler sonucunda elde edilen veriler kullanılarak Tablo 4 hazırlanmıştır.



Şekil 22. 3 W'lık LED sürücüleri **Şekil 23.** 5 W'lık LED sürücüleri

Şekil 22 ve şekil 23'te bazı devre topolojileri gösterilmiştir. Bu devre topolojilerinin amacı daha küçük güçte sürücü devreleri kullanarak daha yüksek güçte çıkış veren devreler tasarlamaktır. Yani toplam çıkış gücü 15 W ve yaklaşık olarak 1200 lm değerinde ışık akısı gerektiren sistem için böyle bir sürücü devre topolojisine ihtiyacımız olduğunda, bunu yukarıda gösterildiği gibi 5 tane 3 W'lık veya 3 tane 5W'lık sürücü devrelerini şekildeki gibi bağlayarak elde edebiliriz.

Bu devre topolojileri enerji kaybı bakımından incelendiğinde Şekil 22'deki devre topolojisi için, 3 W'lık LED sürücü devresinin verilerine bakıldığında (Tablo 4), 24 Volt giriş geriliminde toplam güç kaybı 0.78 W olmaktadır. Bu topolojide biz 5 tane bu sürücüden kullandığımız için $0.78 \times 5 = 3,9$ W' tır. Şekil 23'deki devre topolojisi için, 5 W'lık LED sürücü devresi kullanılan devre topolojisindeki kayıp $3 \times 0.85 = 2,55$ W olduğu görülmektedir.

Devre topolojileri ışık verimliliği açısından incelendiğinde ise, Tablo 4'e göz attığımızda 3 W'lık LED'in yaydığı ışık akısı miktarı 160 lm -190 lm arasında değişirken, $5 \times 190 \text{ lm} = 950$ lm ışık vermektedir buda istediğimiz değerin çok altındadır. Biz bu sistemi 5 W'lık sürücülerle uyguladığımızda, 5 W'lık LED'in yaydığı ışık akısı miktarı 400 lmdir. Devrede toplam 3 tane LED kullanacağımızdan $3 \times 400 = 1200$ lm yapacaktır. Bu değerde kullanmak istediğimiz ortamın gerektirdiği ışık akısı miktarı için yeterli olacaktır.

4. SONUÇ

LED'lerin diğer aydınlatma elemanlarına göre avantajları vardır. Bunlar, LED'lerin uzun ömürlü olması, bakım sürelerinin kısa ve az maliyetli olması, boyutlarının küçük olması ve en önemlisi yüksek enerji verimliliği sağlamasıdır. Bu sebeplerden dolayı LED'ler son 10 yılda ve önümüzdeki yıllarda diğer aydınlatma elemanları ile yer değiştirmeye başlamış ve çok hızlı bir şekilde de değiştirmeye devam edecektir.

Güç LED'leri standart LED'lere göre çok fazla akım çekmektedir. LED'lerin verimli bir ışık kaynağı olarak kullanılabilmesi için belirli bir sabit akımda çalışması gerekmektedir. LED'lere sabit akım sağlama bilmek için sabit akım sınırlayıcı devrelere ihtiyaç vardır. Bu çalışmamızda farklı devre topolojileriyle kullanılacak LED'in ihtiyacı olan akımı sağlayan bazı devreler bilgisayar ortamında tasarlanmış, analizleri yapılmış ve sonuçlar elde edilmiştir.

Tablo 2, 3 ve 4 incelendiğinde uzun süreli kullanımlarda tek dirençle yapılan ve LM117 regülatörü ile yapılan devre topolojilerinin çok verimli olmadığı yaptığımız simülasyonlarda ve hesaplamalarda görülmüştür. LM3406 Entegresi ile yaptığımız DC-DC düşürücü (buck) konverter tabanlı LED sürücülerinin ise daha verimli olduğu gözlenmiştir. Yaptığımız çalışmada elde ettiğimiz sonuçlara göre (Tablo 4) aynı giriş gerilimine sahip farklı güçlerde yapılan LED sürücülerinin verimliliği, sürücü devre gücü ile doğru orantılıdır. Devre gücü arttıkça verimlilikte artmaktadır. 1 W, 3 W ve 5 W'lık sürücü devrelerinin verimlilikleri kıyaslandığında en verimli sürücü devrenin 5 W ile yapılan devre topolojisi olduğu görülmektedir. Ayrıca daha küçük devreleri kullanarak elde ettiğimiz daha yüksek güçteki devre topolojilerinde ise Şekil 23'teki devre topolojisinin, hem kayıplarının az olması yönünden ve hem de ışık akısı yönünden incelendiğinde daha verimli olduğu belirlenmiştir.

5. KAYNAKLAR

- [1]Sermin Onaygil, Aydınlatmada Enerji Verimliliği: LED Teknolojisi, Elektrik Mühendisliği Dergisi Sayı-446 sayfa: 29, Ocak 2013
- [2]Ron LENK, Carol LENK, Practical Lighting Design with LEDs, syf:2, 2011
- [3]Wing Yan Leung, Tsz Yin Man, Mansun Chan, A High-Power-LED Driver with Power-Efficient LED Current Sensing Circuit, 978-1-4244-2362-0/08, 2008 IEEE.
- [4]The_LED_History abctronics, [Online] http://www.abctronics.com/media/abcT_The_LED_History.pdf
- [5]W.D. van Driel X.J. Fan, Solid State Lighting Reliability Components to Systems, 2013, syf: 14
- [6]Humbat Nasibov, Ertan Balaban, Alisher Kholmatov, Adalat Nasibov, High-brightness, high-power LED-based strobe illumination for double-frame micro particle image velocimetry, Flow Measurement and Instrumentation 37; (2014) 12–28
- [7]Patrick Mottier, LEDs for Lighting Applications, 2009, syf: 4
- [8]Aníbal De Almeida, Bruno Santos, Bertoldi Paolo, Michel Quicheron, Solid state lighting review – Potential and challenges in Europe, 2014
- [9]Coaton, J. R (2001). The genesis of incandescent lamp manufacture. Engineering Science And Education. 11, 17-24
- [10]LM117/LM317A/LM317-N Three-Terminal Adjustable Regulator datasheet. [Online] <http://www.ti.com/lit/ds/symlink/lm117.pdf>
- [11]Hasan Komurcugil, Non-singular terminal sliding-mode control of DC–DC buck converters, 2012 Elsevier
- [12] LM3406/06HV1.5A Constant Current Buck Regulator for Driving High Power LEDs [Online] <http://www.ti.com/lit/ds/symlink/lm3406hv-q1.pdf>

Determination Power Correction Parameters of Buck and Buck Boost LED drivers for LED lighting

Ö. F. Farsakoğlu¹, İ. Çelik², İ. Atik³, H. Y. Hasirci⁴

^{1,2,3,4}Electrical and Electronics Engineering Department, Kilis 7 Aralık University, Faculty of Engineering and Architecture, 79000, Kilis, TURKEY

¹ffarsakoglu@kilis.edu.tr, ²ibrahimcelik@kilis.edu.tr, ³ipekinal@kilis.edu.tr,
⁴hyusufhasirci@kilis.edu.tr

ABSTRACT

Today, power LEDs have properties of durability, long-life, and high efficiency. Due to this reason, they are widely used for lighting applications. As LED drivers for lighting applications are connected to the electricity distribution network, they should be complied with power terms. In this case, the harmonic distortion between the AC mains voltage and current and displacement factor can be maintained as low as possible. Today, power supplies used for lighting applications in Europe are determined in accordance with EN / IEC 61000-3-2 standards. These standards include power factors and total harmonic distortions. In recent years, taking into account these standards, power factor correction (PFC) techniques have been developed for LED lighting applications. In this study, the buck-boost converter and buck converter structures were discussed. Simulations of active PFC Buck LED driver at Conduction Mode (CCM) were prepared. Also, Buck-Boost LED drivers were discussed on this way. According to the results of simulations, the data of LED driver Power Factor (PF) and Total Harmonic Distortion (THD) were obtained. The results were examined in accordance EN / IEC 61000-3-2 standards.

Keywords: power LED, power factor correction, EN / IEC 61000-3-2 standards , PFC Buck LED driver, PFC buck-boost LED driver.

1. GİRİŞ

Teknolojik gelişmelere paralel olarak artan enerji ihtiyacının büyük çoğunluğunu elektrik enerjisi oluşturmaktadır. Elektrik enerjisi üretimi kömür, petrol, doğal gaz, su potansiyeli, nükleer yakıtlar, güneş, jeotermal enerji gibi pekçok kaynaktan üretilmesinin yanında bu kaynakların sınırlı olması elektrik enerjisinin verimli kullanılmasını çok daha önemli hale getirmiştir. Elektrik enerjisinin aydınlatmada kullanım payı toplam enerji kullanımına göre azımsanmayacak kadar büyüktür. Dünya genelinde, aydınlatmada kullanılan enerjinin toplam elektrik enerjisine oranı %18 olarak verilmektedir. Tüketimin artması ve elektrik enerji kaynaklarının sınırlı olması nedeniyle diğer enerji sistemlerinde olduğu gibi aydınlatma sistemlerindeki verim de gittikçe önem kazanmaktadır [1].

Aydınlatma sistemlerinde verimli bir elektronik malzeme olarak kullanılan güç LED'lerinin yanı sıra güç LED'lerin sürme devrelerinin de verimli olması oldukça önemlidir. Güç LED yüküne göre sabit akım çıkış verebilmeli ve sunulan standartlara (Aydınlatma ürünlerinde kullanılan güç kaynaklarına yönelik standartlar olan EN / IEC 61000-3-2) [2] uygun olarak tasarlanmalıdır [3]. Bu standartlar göz önüne alınarak son yıllarda LED'li aydınlatma uygulamalarında güç faktörü düzeltme (PFC) teknikleri geliştirilmiştir. Güç faktörü düzeltme teknikleri sayesinde standartlara uygun LED

sürücü tasarımı gerçekleştirilebilmektedir. Ancak güç faktörü düzeltmeli LED sürücülerde toplam maliyeti düşürürken, sürücü performansını optimize etmek önemli hale gelmektedir [4]. Bu çalışmada, ilk olarak güç faktörü tanımı ve IEC 61000 -3-2 standartları ele alınmıştır. Daha sonraki bölümlerde ise buck-boost ile buck dönüştürücü yapısı hakkında bilgiler verilmiş ve buck-boost ile buck temelli güç faktörü düzeltme devreleri incelenmiştir. Son kısımda da buck temelli aktif güç faktörü düzeltme devresi analizi gerçekleştirilmiştir.

2.GÜÇ FAKTÖRÜNÜN TANIMI

2.1 Doğrusal Yük Durumunda (İdeal sinüs dalga formları)

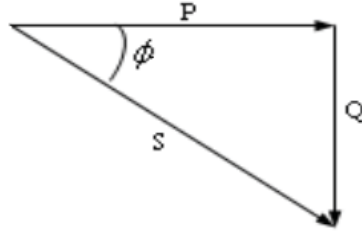
Güç faktörü, PF, AC şebekeye bağlı bir yük tarafından çekilen aktif gücün görünür güce oranı olarak tanımlanır. Aşağıdaki gibi ifade edilir.

$$PF = \frac{P}{S} \quad (1)$$

Klasik güç faktörü ifadesi, AC şebeke gerilimi ile akımının ideal sinüsoidal olması durumunda aşağıdaki gibi olur.

$$PF = \cos \phi \quad (2)$$

Burada $\cos \phi$ akım ile gerilim arasındaki yerdeğiştirme faktörünü ifade etmektedir (Şekil 1). Klasik anlamda PFC'nin anlamı yerdeğiştirme faktörünü kompanse etmektir.



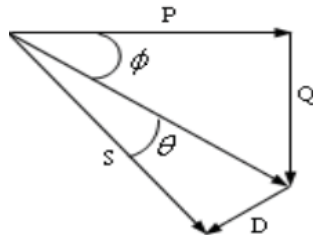
Şekil 1. Doğrusal yüklerde güç faktörü[5]

2.2 Doğrusal Olmayan Yük Durumunda

Nonlineer yüklerde AC şebeke akımı sinüsoidal olmayacaktır. Bu nedenle klasik güç faktörü ifadesi geçerliliğini yitirecektir. Güç faktörü ifadesi sinüsoidal gerilim ve sinüsoidal olmayan akım için aşağıdaki gibidir.

$$PF = \frac{V_{rms} \cdot I_{rms1} \cdot \cos \phi}{V_{rms} \cdot I_{rms}} = \frac{I_{rms1} \cdot \cos \phi}{I_{rms}} = K_d \cdot \cos \phi \quad (3)$$

Burada bozulma faktörü olarak tanımlanan $K_d = \cos \theta = I_{rms1} / I_{rms}$ 'dir. Bu durumda güç faktörü bozulma faktörü ile yer değıştirme faktörüne bağlıdır [6].



Şekil 2. Doğrusal Olmayan yüklerde güç faktörü [5].

Güç faktörü düzeltilmeli AC güç kaynaklarında bozulma faktörü, K_d ile giriş akım toplam harmonik bozulması, THD arasındaki bağıntı aşağıdaki gibi ifade edilebilir.

$$Kd = \frac{1}{\sqrt{1 + THD^2}} \quad (4)$$

IEC 61000-3-2 standartlarına göre giriş akım toplam harmonik bozulması aşağıdaki gibi ifade edilir [7].

$$THD = \frac{\sqrt{\sum_{n=2}^{40} I_n^2}}{I_1} \quad (5)$$

Burada I_1 ve I_n 1. harmonikten n. harmoniğe kadar akımların ayrı ayrı genlikleri ifade eder.

Buradan PF için tanım, yerdeğiştirme faktörü ve bozulma faktörü hesaplanarak aşağıdaki gibi ifade edilir.

$$PF = \frac{\cos \phi}{\sqrt{1 + THD^2}} \quad (6)$$

Bu ifadede $\cos \phi = 1$ olduğu durumda yani akım ile gerilim arasındaki yerdeğiştirme faktörü olmadığında yüksek güç faktörü ve düşük THD her ikisi birden elde edilebilir.

3. IEC 61000 -3-2 standartları

Güç LED sürücülerinin AC şebekeden beslendikleri için AC şebekedeki Toplam Harmonik Bozulmaların ve akım ve gerilim arasındaki faz farkının olabildiğince küçük olması istenmektedir. Bu sayede kaliteli bir güç tüketimi gerçekleşmiş olacaktır. Günümüzde, Avrupa'da aydınlatma uygulamalarında kullanılan güç kaynaklarının güç kalitesi, güç faktörü ve harmonik içerik için IEC-61000-3-2 standardı kullanılmaktadır. Japonya'da ise bunlar JIS 61000-3-2 standardıyla belirlenmektedir [8]. Bunların yanında Energy Star gibi ABD Enerji Departmanı ve ABD Çevre Koruma Ajansı tarafından oluşturulmuş programlarda ticari uygulamalar için güç faktörünün en az 0.9, mesken uygulamaları içinse en az 0.7 olması gerekmektedir [8].

Uluslararası Enerji (IEC) Komisyonu tarafından [IEC 61000 -3-2 Standardı (2005) ve eşdeğeri olan EN 61000 -3-2 Standardı (2006)] standartlarında elektromanyetik uyumluluğa değinilmektedir. Aynı zamanda sözkonusu standartlarda AC şebekede izin verilen maksimum harmonik akımların değerleri belirtilmektedir. Ülkemizde ise bu standart TS EN 61000 -3-2 adı ile uygulanmakta olup alçak gerilimde faz başına 16 A'den düşük akım çeken cihazların harmonik akım sınırlarını belirlemektedir [10].

2.1. C Sınıfı Donanımlar için Sınır Değerler

Aydınlatma cihazları 25W'dan daha fazla aktif güce sahip olan C sınıfı cihazlar için sınır Tablo 1 ve Tablo 2'de verilmiştir. Maksimum izin verilen harmonik akımlar temel giriş akımının yüzdesi olarak ifade edilmiştir. Giriş gücü 25W'a eşit veya daha küçük C sınıfı teçhizat için aşağıdaki şartlardan biri sağlanmalıdır;

- Tablo 2 (sütun iki) sınırları geçerlidir.
- Üçüncü harmonik akım temel akımın % 86 geçemez ve beşinci harmonik akım % 61 geçemez.[10]

Tablo 1. C sınıfı cihazların için sınırlamaları

| Harmonik Düzenlemeler N | C Sınıfı | |
|-------------------------------|---|--|
| | Maksimum izin verilen harmonik akımın temel frekansta giriş akımın yüzdesi olarak ifadesi (%) | |
| Tek harmonikler | | |
| 3 | 30×λ* | |
| 5 | 10 | |
| 7 | 7 | |
| 9 | 5 | |
| 11 | 3 | |
| 13 | 3 | |
| 15≤n≤39 | 3 | |
| Çift Harmonikler | | |
| 2 | 2 | |
| 4 | - | |
| 6 | - | |
| 8≤n≤40 | - | |
| λ* : devre güç faktörü | | |

Tablo 2. D sınıfı cihazlar için sınırlamalar

| Harmonik Düzenlemeler n | Sınıf D | |
|-------------------------------|---|--|
| | Maksimum izin verilen harmonik akım watt başına mA/W | Maksimum izin verilen harmonik akım A |
| Tek harmonikler | | |
| 3 | 3.4 | 2.3 |
| 5 | 1.9 | 1.14 |
| 7 | 1.0 | 0.77 |
| 9 | 0.5 | 0.40 |
| 11 | 0.35 | 0.33 |
| 13 | 3.85/n | 0.21 |
| 15≤n≤39 | 3.85/n | 0.23×8/n |

4.Buck (Düşürücü) Dönüştürücü

En yaygın kullanılan anahtarlamalı dönüştürücü buck tipidir. Girişine uygulanan gerilimleri aynı polaritede daha düşük seviyede bir çıkış gerilimine çevirirler. Şekil 3(a)'da buck dönüştürücü devre şeması verilmiştir. Buck dönüştürücü de kullandığımız anahtar, dönüştürücünün kaynak gerilimiyle bobin arasındaki bağlantıyı açıp kapatmaktadır [4]. Anahtarın iletimde olduğu süre boyunca kaynak gerilimi yük uçlarına bağlıdır ve güç akışı kaynaktan yüke doğru gerçekleşmektedir. Anahtarın kesimde olduğu süre boyunca ise yük akımı serbest diyot üzerinden akışını tamamlar [11].

Şekil 3(b)'de düşürücü dönüştürücünün bobin gerilim ve akımı dalga şekilleri verilmiştir. Anahtar kesimde olduğu durumda bobin üzerindeki gerilim ifadesi;

$$V_L = V_S - V_0 = L \cdot \frac{di_L}{dt} \quad (7)$$

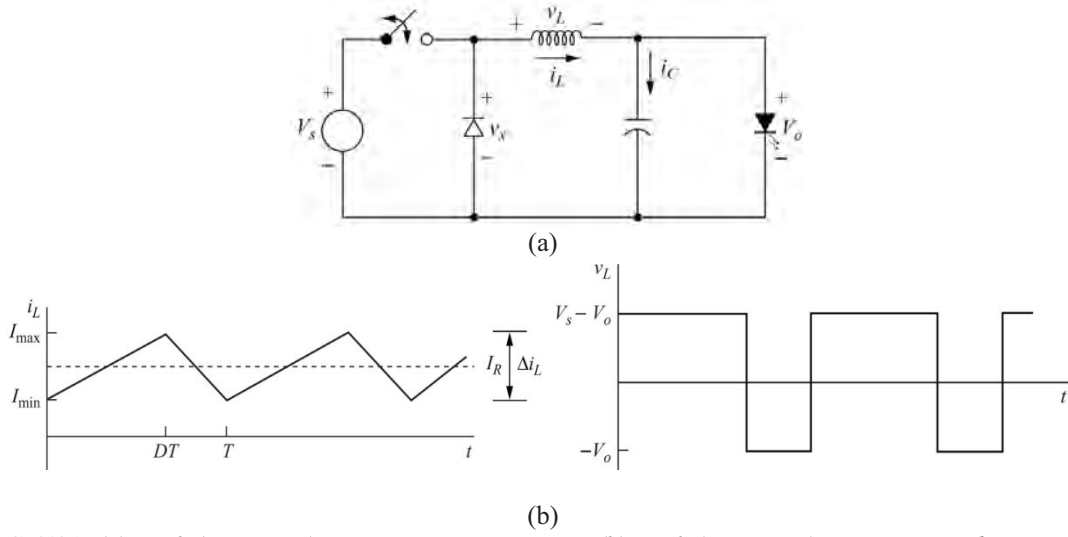
bağıntısı ile verilir. Anahtar iletimde olduğu durumda bobin üzerindeki gerilim ifadesi ise;

$$V_L = -V_0 = L \cdot \frac{di_L}{dt} \quad (8)$$

bağıntısı ile verilmektedir. Ortalama çıkış gerilimi ifadesi;

$$V_0 = V_S \cdot D \quad (9)$$

bağıntısını sağlamaktadır.



Şekil 3. (a) Buck (Düşürücü) Dönüştürücü Devresi ve (b) Buck (Düşürücü) Dönüştürücü devresinin bobin gerilim ve akımı dalga şekilleri[12]

5. Buck-Boost (Düşürücü-Yükseltici) Dönüştürücü

Buck-Boost (Düşürücü-Yükseltici) yapısında giriş geriliminin genlik ve polaritesini değiştirmek mümkündür. Şekil 4(a)'da Buck-Boost (Düşürücü-Yükseltici) dönüştürücü yapısı görülmektedir. Anahtarın iletimde olduğu süre boyunca diyot ters polaritede olup, endüktans üzerinde enerji depolanır. Anahtar kesimde olduğu sürede, diyot doğru polaritede olup endüktans üzerindeki enerji çıkışa aktarılır. Şekil 4(b)'de düşürücü dönüştürücünün bobin gerilim ve akımı dalga şekilleri verilmiştir. Anahtar kesimde olduğu durumda bobin üzerindeki gerilim,

$$V_L = V_S = L \cdot \frac{di_L}{dt} \quad (10)$$

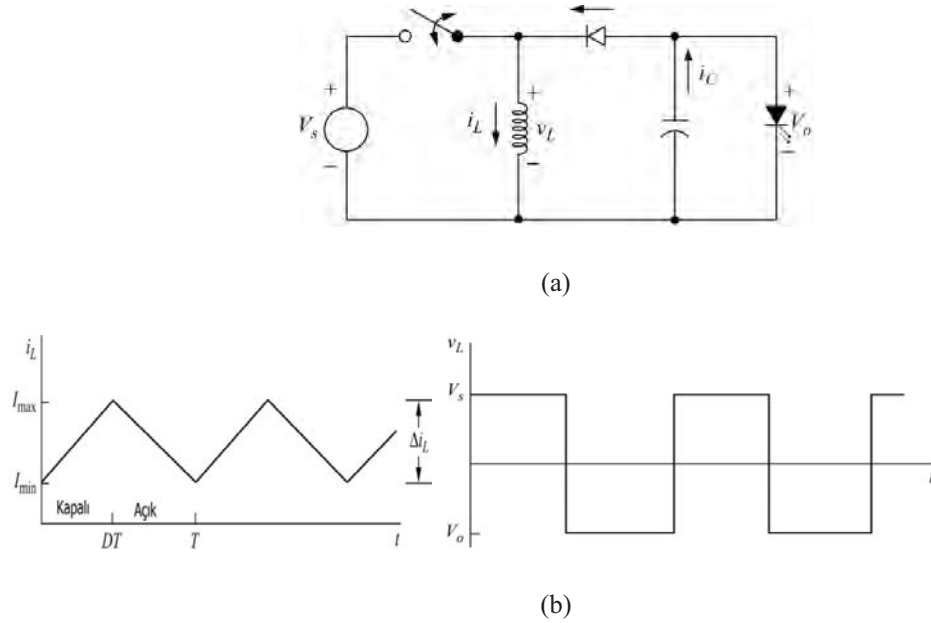
anahtar iletimde olduğu durumda bobin üzerindeki gerilim,

$$V_L = V_0 = L \cdot \frac{di_L}{dt} \quad (11)$$

Ve ortalama çıkış gerilimi,

$$V_o = -V_s \cdot \frac{D}{1-D} \quad (12)$$

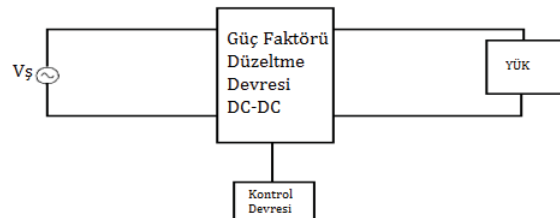
bağıntılarıyla verilir.



Şekil 4. (a) Buck-Boost (Düşürücü-Yükseltici Dönüştürücü Devresi ve (b) Buck-Boost (Düşürücü-Yükseltici Dönüştürücü devresinin bobin gerilim ve akımı dalga şekilleri[12]

6. Tek Aşamalı Güç Faktörü Düzeltme Devre Yapısı

Tek aşamalı bir GFD sisteminin genel blok şeması Şekil 5'te görülmektedir. Tek aşamalı güç faktörü düzeltme devrelerinde, çıkış gerilimi regülasyonu ve güç faktörü düzeltme işlemleri tek bir kontrolör tarafından gerçekleştirilmektedir. Bu devre yapısı basit ve az elemanla oluşturulabilmektedir. Bunun yanında devrenin veriminin artırılması ve elemanların boyutlarının azaltılması bakımından, gücün çıkışa tek seferde işlendiği tek aşamalı güç faktörü düzeltme devrelerine olan ilgi gittikçe artmaktadır [13]. Halihazırda geliştirilen özel tümeleşik devre elemanları sayesinde tek aşamalı güç faktörü düzeltme metodu yaygın olarak kullanılmaya başlanmıştır. Literatürde bu konuda çok sayıda topoloji ortaya konulmuş, bu topolojilerde doğrudan aktarılan gücün oranının artırılması hedeflenmiştir.

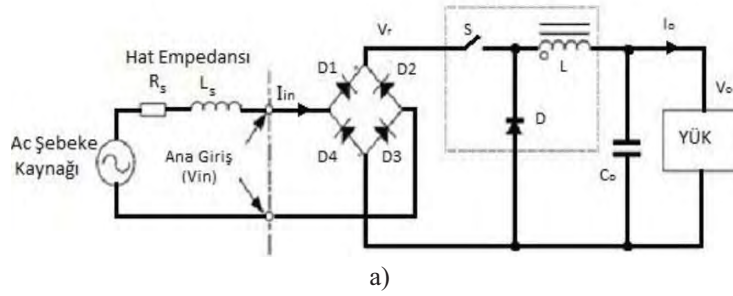


Şekil 5. Tek aşamalı güç faktörü düzeltme devreleri yapısı [14]

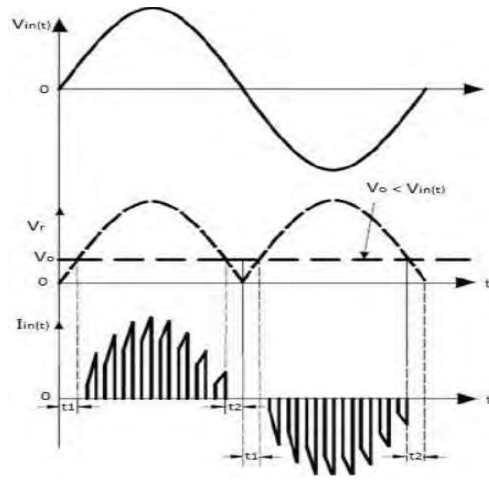
6.1 Buck (Düşürücü) Temelli Dönüştürücü Yüksek frekans aktif Güç Faktörü Düzeltme

Şekil 6(a) ve (b)'de giriş gerilimini azaltan dönüştürücü temelli GFD devresi ve dalga şekilleri gösterilmiştir. Dönüştürücü giriş gerilimi $V_{in}(t)$, çıkış gerilimini V_o 'dan yüksek olduğu durumda çalışmaktadır. AC girişten t_1 ile t_2 süreleri boyunca akım akmamaktadır. Sıfır geçiş giriş gerilim yakınında şebeke akım zarfında bozulma meydana getirir. Bunun yanında, bobin akımı kesintisiz olsa

da dönüştürücünün giriş anahtarlama akımı S anahtarı her anahtarlama periyodunda girişe etki ettiği için kesintilidir. Bu sebeple giriş akımı önemli yüksek frekans bileşenlere sahiptir ve bu EMI girişimlerinin ve filtreleme gereksinimlerinin artmasına neden olmaktadır [15].



a)

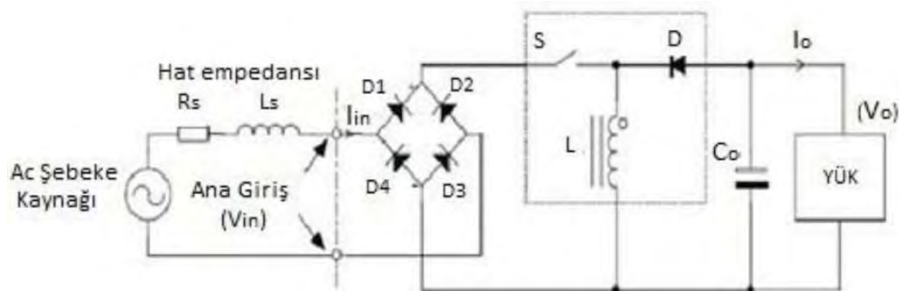


b)

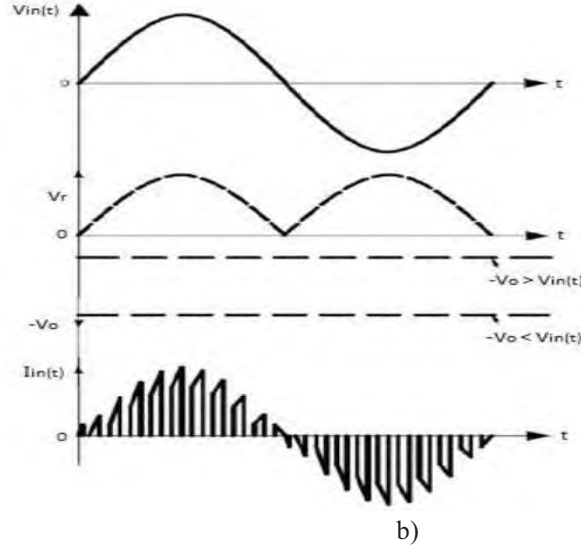
Şekil 6. Yüksek frekans Aktif GFD; (a) Buck (Düşürücü) Dönüştürücü temelli Devresi ve (b) Buck (Düşürücü) Dönüştürücü temelli devresinin gerilim ve akım şekli [17].

6.2. Buck-Boost (Düşürücü-Yükseltici) Temelli Aktif Güç Faktörü Düzeltme

Şekil 7(a)'da ve Şekil 7(b)'de ise buck - boost dönüştürücü temelli GFD devresi ve dalga şekilleri gösterilmiştir. Bu devre giriş akımını hem yükseltebilir hem de düşürülebilir. Çıkış geriliminin tersine çevrilmesinden dolayı anahtar üzerinde gerilim stresi oluşmaktadır. Dönüştürücünün şebeke periyodunda çalışabildiği durumda giriş akımında geçiş bozulmalarına rastlanmaz. Fakat buck dönüştürücüde olduğu gibi bobin akımı kesintisiz ise dönüştürücünün giriş anahtarlama akımı kesintilidir çünkü anahtar giriş akımına etki eder. Bu yüzden giriş akımı elektro manyetik girişimlerini ve filtre gereksinimlerini arttıran yüksek frekanslı içeriklere sahiptir [16].



Şekil 7. (a) Yüksek frekans Aktif GFD; Buck-Boost (Düşürücü-Yükseltici) Dönüştürücü temelli Devresi



Şekil 7. (b) Yüksek frekans Aktif GFD; Buck-Boost (Düşürücü-Yükseltici) Dönüştürücü temelli devresinin gerilim ve akım şekli[17]

6.2.1. Buck-Boost (Düşürücü-Yükseltici) Temelli Aktif Güç Faktörü Düzeltme Devresi Literatür Çalışması

Bu bölümde IEC 61000-3-2 standartlarına uygun olarak gerçekleştirilmiş bir PFC Buck Boost devresinin deneysel performansı sonuçları Tablo 3'te verilmiştir. LED sürücü tasarımında ileri yön gerilimi 3.1 V, ileri yön akımı 325mA olan 1W'lık 13 tane LED seri bağlanmış ve anahtarlama frekansı 66kHz'dir. PFC Buck-Boost sürücü devresinin verimliliği %80.81-%83.41 arasında değişmektedir. AC giriş akımındaki toplam harmonik bozulma % THD değeri %12.24-%20.46 arasında olduğu Tablo 3'te görülmektedir.

AC giriş gerilimi V_{in} değerinin 90 V ile 270 V arasında artış gösterimiyle birlikte giriş akım toplam harmonik bozulma, % THD değerinin'de 12.24'ten 20.46'ya yükseldiği görülmektedir. Çıkış yükünün sabit olduğu (13W) durumlarda en yüksek toplam harmonik bozulma en yüksek giriş geriliminde meydana gelmektedir. Bunun yanında düşük çıkış yüküne rağmen PFC Buck Boost devresinin çok yüksek verimlilikte çalıştığı gözlenmektedir.

PFC Buck Boost devresinin verimiyle güç faktörü arasında doğrudan bir ilişki kurmak mümkün olmasa da yüksek güç faktörüne sahip bir sürücünün veriminde yüksek olacağını söylemek mümkündür. Bu devre yapısı standartlara uygunluğunun yanı sıra yüksek veriminden dolayı düşük güçler için ideal bir yapıdır.

Tablo 3. Buck Boost (Düşürücü-Yükseltici) temelli tek katlı LED sürücü performans indisleri [18]

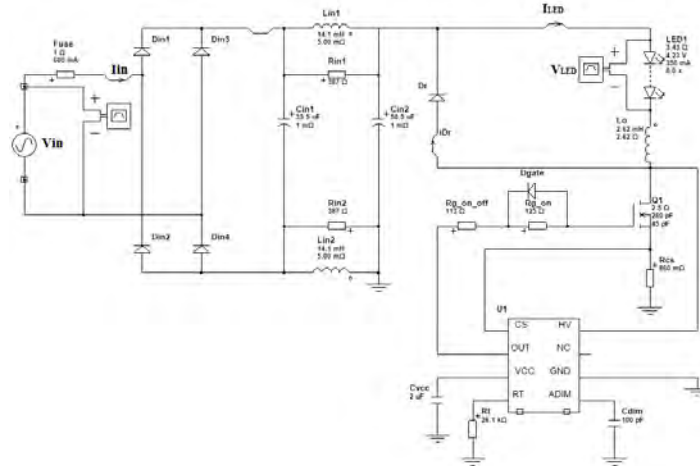
| V_{in} (V) | I_{in} (mA) | V_{LED} (V) | I_{LED} (mA) | P_{out} (W) | P_{in} (W) | Verim(%) | PF | (%) THD |
|--------------|---------------|---------------|----------------|---------------|--------------|----------|-------|---------|
| 90 | 170.51 | 40.553 | 304.3 | 12.34 | 15.27 | 80.81 | 0.992 | 12.24 |
| 130 | 121.14 | 40.543 | 317.7 | 12.88 | 15.60 | 82.56 | 0.987 | 15.93 |
| 180 | 87.55 | 40.411 | 319.8 | 12.92 | 15.48 | 83.46 | 0.982 | 18.82 |
| 220 | 70.56 | 40.265 | 314.7 | 12.67 | 15.19 | 83.41 | 0.978 | 19.96 |
| 270 | 58.71 | 40.195 | 307.8 | 12.37 | 14.87 | 83.18 | 0.973 | 20.46 |

6.3.1 Buck (Düşürücü) Temelli Dönüştürücü Yüksek Frekans Aktif Güç Faktörü Düzeltme Devresinin Uygulaması

Şekil 8’de Buck temelli tek katlı LED sürücü devresi gösterilmiştir. Bu sürücüde kontrol işlemi FL7701 entegresi tarafından gerçekleştirilmektedir. FL7701 entegresi Buck konvertör topolojisi Sürekli iletim modu (CCM) için bir dijital kontrol algoritması kullanılarak akıllı PFC fonksiyonu ile temel bir PWM kontrolördür. FL7701 entegresi otomatik olarak giriş gerilim şartını algılar ve yüksek güç faktörü için dahili referans sinyali gönderir. FL7701 entegresine AC giriş uygulandığında PFC fonksiyonu otomatik olarak seçiliyor. Aksi takdirde, yani DC giriş uygulandığında PFC fonksiyonu otomatik olarak devre dışı bırakılıyor [19].

Bu çalışmada kullanılan PFC Buck devre sürücü simülasyonu Power Supply WebDesigner’da gerçekleştirilmiştir. PFC Buck temelli tek katlı LED sürücü devresinde 8 adet 1 W’lık Power LED yük olarak kullanılmıştır. Power LED’lerimiz ileri yön gerilimi 3.8 V ve ileri yön akımı 350 mA’dır. PFC Buck temelli sürücüye uygulanan AC giriş gerilimi 130 V ve 310 V değerleri arasında değişmektedir. Anahtarlama frekansı ise 66 kHz’dır. AC giriş gerilimi 130 V ve 310 V değerleri arasında belirlenmiş giriş gerilimlerinde PFC Buck LED sürücünün giriş gücü (P_{in}), giriş akımı (I_{in}), çıkış gücü (P_{out}), çıkış akımı (I_{LED}), çıkış gerilimi (V_{LED}), güç faktörü (PF), bozulma faktörü (DF) ve LED sürücünün verimliliği Tablo 4’te verilmiştir. Bunun yanında Şekil 9(a)-(c)’de PFC Buck LED sürücüye uygulanan AC giriş gerilim ve akım değerleri 130V, 210V ve 311V için gösterilmiştir. Ayrıca Şekil 10(a)-(c)’de AC giriş geriliminin 130V, 210V ve 311V değerlerinde toplam harmonik bozulma değerleri ifade edilmiştir.

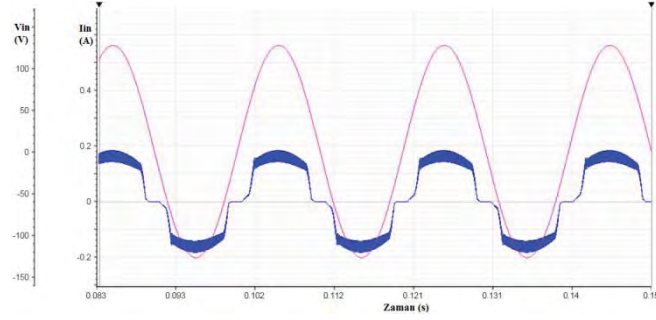
AC giriş gerilim değeri 130 V ile 311 V arasında değiştiğinde, PFC Buck LED sürücü devresinin verimliliği % 79.80- % 83.78 arasında değişmektedir. AC giriş akımındaki toplam harmonik bozulma % THD değeri % 30.03- % 38.87 arasında değişmektedir. Simülasyonu gerçekleştirilen PFC Buck LED sürücüsü IEC 61000-3-2 Sınıf C standartlarında belirtilen 25 W altında aydınlatma uygulamaları için sınır değerler içerisindedir. Düşük güçlü aydınlatma uygulamalarında yüksek verimliliklerde çalışmaktadır.



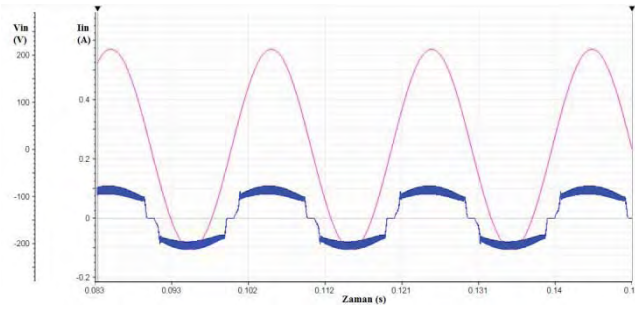
Şekil 8. PFC Buck (Düşürücü) LED sürücü devresi

Tablo 4. PFC Buck (Düşürücü) LED sürücü performans parametreleri

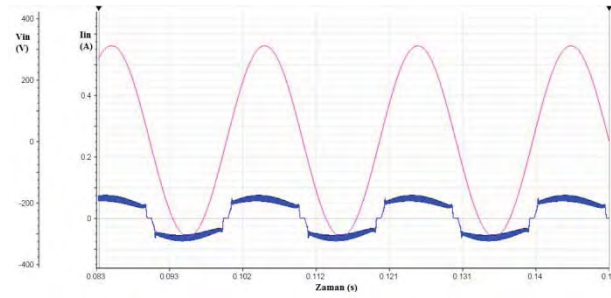
| V_{inrms} (V) | I_{inrms} (mA) | V_{LED} (V) | I_{LED} (mA) | P_{out} (W) | P_{in} (W) | Verim(%) | PF | % THD | DF |
|-----------------|------------------|---------------|----------------|---------------|--------------|----------|-------|-------|-------|
| 90.92 | 125 | 31.38 | 289 | 9.07 | 11.36 | 79.8 | 0.957 | 30.03 | 0.958 |
| 121.24 | 95 | 31.83 | 300 | 9.55 | 11.52 | 82.9 | 0.948 | 31.41 | 0.954 |
| 151.53 | 77 | 32 | 304 | 9.73 | 11.67 | 83.38 | 0.946 | 33.95 | 0.947 |
| 181.86 | 65 | 32.1 | 306 | 9.82 | 11.82 | 83.1 | 0.936 | 36.34 | 0.940 |
| 222.28 | 53 | 32.18 | 307 | 9.87 | 11.78 | 83.79 | 0.928 | 38.87 | 0.932 |



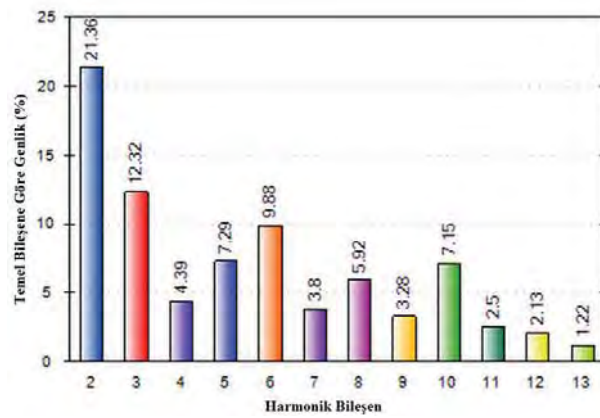
Şekil 9. (a) PFC LED sürücüsüne uygulanan AC gerilim (V_{in}) ve akım değeri (I_{in}) ($V_{inrms}=90.91$)



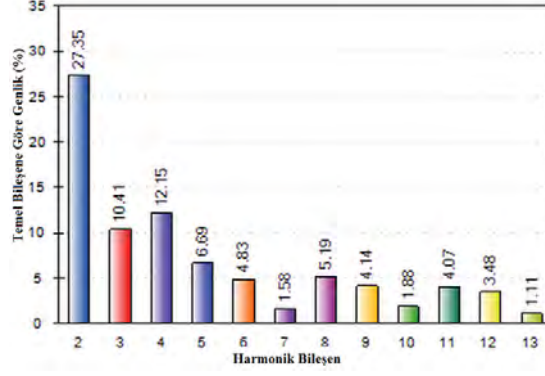
Şekil 9. (b) PFC LED sürücüsüne uygulanan AC gerilim (V_{in}) ve akım değeri (I_{in}) ($V_{inrms}=151.55$)



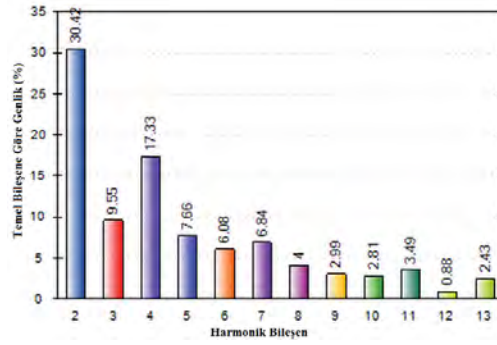
Şekil 9. (c) PFC LED sürücüsüne uygulanan AC gerilim (V_{in}) ve akım değeri (I_{in}) ($V_{inrms}=222.28$)



Şekil 10. (a) PFC LED sürücüsüne 130V AC giriş gerilim uygulandığında harmonik bileşen spektrumu



Şekil 10. (b) PFC LED sürücüsüne 210V AC giriş gerilim uygulandığında harmonik bileşen spektrumu



Şekil 10. (c) PFC LED sürücüsüne 311V AC giriş gerilim uygulandığında harmonik bileşen spektrumu

7. SONUÇ

LED aydınlatma uygulamalarında sistemin kalitesi, verimliliği ve uzun ömürlü olması LED sürücü tasarımına bağlıdır. Bu çalışmada PFC Buck LED sürücü devresi tasarımındaki amaç AC giriş gerilim değerleri 130V - 310V arasında değiştiğinde sistemin güç kalitesini ve verimliliğini arttırmaktır.

Bu çalışmada düşük güçlerde kullanılan PFC Buck LED sürücü devresi ve Buck Boost sürücü devreleri incelenmiştir. Buck Boost tabanlı LED sürücü devreleri için sadece literatür taraması yapılmıştır. Bu literatür taramasında Buck Boost PFC LED sürücünün giriş gücü (P_{in}), giriş akımı (I_{in}), çıkış gücü (P_{out}), çıkış akımı (I_{LED}), çıkış gerilimi (V_{LED}), güç faktörü (PF) ve LED sürücünün verimliliği ele alınmıştır. AC giriş gerilim değeri 90 V ile 270 V arasında değiştiğinde, toplam harmonik bozulma değeri giriş geriliminin artmasıyla arttığı görülmüştür. Tasarımı yapılan PFC Buck tabanlı LED sürücü devresi düşük güçlü güç LED uygulaması için ele alınmıştır. Gerçekleştirilmesi basit olan bu LED sürücü devresinde bazı değerler [giriş gücü (P_{in}), giriş akımı (I_{in}), çıkış gücü (P_{out}), çıkış akımı (I_{LED}), çıkış gerilimi (V_{LED}), güç faktörü (PF) ve LED sürücünün verimlilik] elde edilmiştir. Devrenin toplam harmonik bozulması ise IEC 61000-3-2 standartlarına göre hesaplanmıştır. AC giriş gerilimi 130 V olduğunda toplam harmonik bozulma değerleri %30.03, verim %79.8 ve güç faktörü 0.957'dir. AC giriş gerilimi 311 V olduğunda ise değeri % 38.87, verim %83.79 ve güç faktörü 0.928 olduğu Tablo 4'te görülmektedir.

IEC 61000-3-2 standartları C sınıfı 25 W altında standartlara göre değerlendirildiğinde Buck temelli sürücünün uygunluğu görülmüştür. Bunun yanı sıra AC giriş gerilimi yüksek olduğu düşük güçlü LED sürücü devrelerinde Buck temelli LED sürücü devresinin verimliliği Buck Boost temelli LED sürücü devrelerine göre yüksek olduğu belirlenmiştir. Fakat literatür taraması yapılan Buck Boost temelli LED sürücü devresi, Buck (Düşürücü) temelli LED sürücüsüne göre toplam harmonik bozulması daha düşüktür. LED sürücülerin tasarımı uluslararası standartlara uygun şekilde gerçekleştirilmesi durumunda düşük güç seviyelerinde dahi LED sürücünün verimliliğinin yüksek olduğu gözlenmiştir.

8. KAYNAKLAR

- [1] H Yılmaz, A. M. Hava, “ G-Led Uygulamaları İçin Güç Katsayısını Düzeltici Tek Katlı AC-DC Dönüştürücülerin İncelenmesi ” EVK 2013, 4. Enerji Verimliliği ve Kalitesi Sempozyumu, Kocaeli, 2011, pp. 144-148.
- [2] Electromagnetic Compatibility (EMC), Part 3-2: Limits –Limits for harmonic current emissions (equipment input current ≤ 16 A per phase), International Standard IEC 61000-3-2, 2005.
- [3] H. Yılmaz, “Tek Katlı Flyback Ve Sepic Gkd’lerin Güç LED’iyle Aydınlatma Uygulaması İçin Tasarımı Gerçeklenmesi Ve Karşılaştırılması”, Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi Fen Bilimleri Enstitüsü, 2012.
- [4] Y.Gürbüz, “Güç Faktörü Düzeltmeli Güç LED Sürücü Tasarımı Ve Gerçekleştirilmesi”, Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, 2012.
- [5] A. Benaissa, M. K. Fellah, A. Massoum "Harmonic reduction based on active solutions," Electronic Journal (Technical Acoustics), Sidi Bel Abbes University, Algeria, 2005. Available at:<http://ejta.org/en/benaissa1>.
- [6] V. Grigore, “Topological Issues In Single-Phase Power Factor Correction”, Helsinki University Of Technology, Finland, 2001.
- [7] Electromagnetic Compatibility (EMC), Part 3-2: Limits –Limits for harmonic current emissions (equipment input current ≤ 16 A per phase), International Standard IEC 61000-3-2, 2005.
- [8] D. G. Vaquero, “Off-Line Supply of Solid-State Lamps. Lamp Modelling, Application of The İntegrated Buck-Flayback Converter and Propesel of a New Optimised Dimming Scheme” University of Oviedo, Spain , 2013.
- [9] Energy Star Program Requirements for Luminaires – Partner Commitments, version 1.1. US Environmental Protection Agency, 2012.
- [10] Harmonic Current Emissions,Guidelines to the standard EN 61000-3-2, [Online] http://www.epsma.org/pdf/PFC%20Guide_November%202010.pdf
- [11] Mert T, “ Güç Faktörü Düzeltme Yöntemlerinin İncelenmesi Ve Bir Uygulama Devresinin Gerçekleştirilmesi”, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, 2007.
- [12] Hart D. W., Power Electronics, *Pearson Education, Inc.*, 2010.
- [13] Y. Jiang, F. C. Lee, G.Hua, W.Tang, “A Novel Single-Phase Power Factor Correction Scheme”, in Applied Power Electronics Conference and Expositin, APEC, 1993, pp. 287-292.
- [14] H. Bodur, E.Akboy, İ. Aksoy “Tek Aşamalı Güç Faktörü Düzeltme Devrelerinin İncelenmesi”, Elektrik- Elektronik ve Bilgisayar Sempozyumu, 2011 , pp. 168-172.
- [15] Kyyra, J. ve Grigore, V., A Step-Down Converter with Low Ripple Input Curret for Power Factor Correction, Helsinki University of Technology, Finland, 2000.
- [16] A. M. Karakaya, “Aktif Güç Faktörü Düzeltme Devrelerinin İncelenmesi Ve Bir Uygulama Devresinin Gerçekleştirilmesi”, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, 2012.
- [17] S. Basu, “Single Phase Active Power Factor Correction Converters”, Chalmers University of Technology, 2006.
- [18] B. Singh, A. Shrivastava, A. Chandra, K. Al-Haddad, “ A Single State Optocoupler-Less Buck-Boost PFC Driver For LED Lamp at Universal AC Mains” 2013,IEEE
- [19] FL7701 Smart LED Lamp Driver IC with PFC Function (Online) <https://www.fairchildsemi.com/datasheets/FL/FL7701.pdf>

Performance Simulation of Gossip Relay Protocol in Multi-hop Wireless Networks

A. Sarı¹, E. Çağlar²

¹Dept. of Management Information Systems

European University of Lefke

Lefke, North Cyprus

²Dept. of Computer Engineering

European University of Lefke

Lefke, North Cyprus

asari@ieee.org, erscaglar@gmail.com

ABSTRACT

The Gossip-Based Relay Protocol (GRP) is developed based on Ad Hoc On Demand Distance Vector Protocol (AODV) and proposed to increase the efficiency of package routing functionality in ad hoc networks through specific flooding scheme. This lightweight protocol reduced the collisions on the network. Request to Send / Clear to Send (RTS/CTS) mechanism is widely used in ad hoc environment with Temporarily Ordered Routing Algorithm (TORA) in order to eliminate collisions and allow access to the shared medium through proposed authentication methods. In this paper, TORA protocol with RTS/CTS mechanism is simulated and compared with proposed GRP in terms of specific performance metrics such as network throughput, end-to-end delay and message flooding rate over the network through OPNET simulation package in order to expose the optimal solution to increase overall network throughput in ad hoc environment.

Keywords – GRP, TORA, RTS/CTS mechanism, OPNET simulation, throughput.

1. INTRODUCTION

Dynamically calculated routing tables and exchange of dynamic routing information among mobile nodes is inevitable for existence of multi-hop networks due to their nature characteristics. The ad-hoc environment provides an opportunity to deploy a simple network without requesting any pre-identified network infrastructure. The lightweight wireless devices which participate in these networks relay the user information on the network and do not generate any dynamic routing information to communicate with other devices. Authentication is prominent for an ad-hoc environment since dynamic routing information is crucial for secure data transfer among participating nodes. On the other hand, due to mobility and dynamic nature of multi-hop wireless networks, participating nodes may not reliable to share static routing information and generation of dynamic routing information becomes a certain issue.

Researchers have discussed Gossip Relay protocols in the literature however there is no detailed research available to evaluate security aspects of GRP since it is designed to enhance performance. The implementation of the GRP relays on the probabilistic method to determine the priority of the message to broadcast by assigning a rate of probability. RTS/CTS mechanism with TORA protocol is used to provide an authentication between sender and receiver so it also creates dynamic routing information with different proposed methods similar to GRP where it aims to decrease the broadcast messages on the network without requesting any explicit route setup. In the next section, all these mechanisms are shown and explained in details and each of the proposed method is simulated in OPNET environment to compare and expose the optimal algorithm or method to use in multi-hop wireless networks.

2. LITERATURE REVIEW

GRP is proposed by the researchers in the literature with an idea of reducing broadcasted messages by a probabilistic method and determine whether a participating node will broadcast the received message or not [1]. The implementation of GRP is designed based on AODV protocol by the researchers which require route calculation and choose the best route to use and forward the messages. Other researchers have proposed another routing efficient algorithm by clustering the participating nodes into subset of neighbors which is efficient and nodes used to forward broadcasted message to the corresponding neighbors rather than sending to entire network [2-3]. The GRP protocol relay on a deterministic approach with probabilistic function in order to reduce message flooding on the network to prevent overhead problem. Based on the GRP protocol, the MAC protocol is modified for CSMA/CA (Collision Avoidance) feature to implement such mechanism on the network. The proposed GRP protocol is used to conduct Parent-Sibling-Child (PSC) relationship for direct traffic flows which clearly shown in the literature [4]. The PSC relationship leads participating nodes to be clustered into such hierarchy to forward broadcasted received message only to it's corresponding relative, eg. child or sibling. The Figure 1, Figure 2 and Figure 3 illustrate the relationship between Parent-Child-Sibling techniques below. Each of the relationship illustrated on the figures have a specific probability formulas which were shown on the formulas below. The P_n represents the broadcast probability where n is received packages. The increased probability for packets received from parent is;

$$P_n = P_{n-1} + (1 - P_{n-1}) * 2^{-3}$$

Where decreased probability for packets received from sibling

$$P_n = P_{n-1} * (1 - (2^{-4}))$$

where n is the nth new message package and n-1 is the n-1th message package.

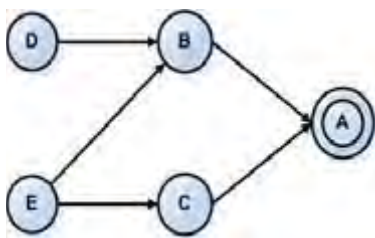


Figure 1. Parent Diagram

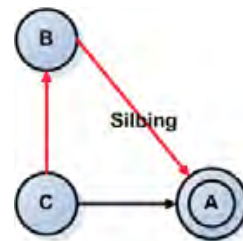


Figure 2. Sibling Diagram

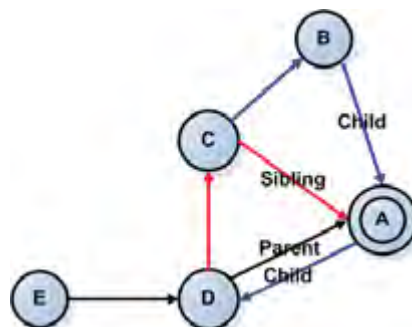


Figure 3. Child Diagram

The RTS/CTS mechanism is a handshaking process that minimizes the problem of collisions when hidden node problem arises on the network. The Figure 4 illustrates the RTS/CTS working mechanism briefly.

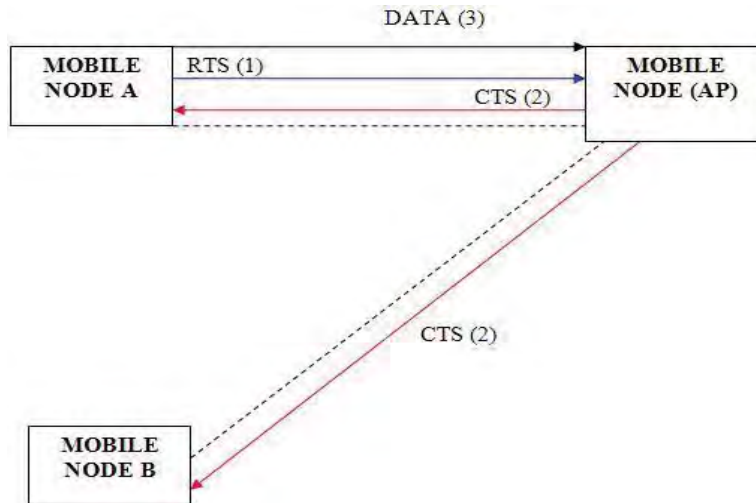


Figure 4. *RTS/CTS Mechanism*

Request to Send (RTS), Clear to Send (CTS) and Fragmentation Thresholds impact on AODV protocol are well known by the simulation experiments conducted in this study and RTS implementation is done in order to increase the performance of the network. In other cases, RTS/CTS is the optional mechanism used by the 802.11 wireless networking protocols to reduce frame collisions introduced by the hidden node problems [5]. The impact of RTS/CTS mechanism is clearly stated by the researchers and exposed through simulation experiment in the literature. However, in order to analyze the difference between the GRP and RTS/CTS under TORA protocol and expose the most optimal way solution for multi-hop wireless networks, it is necessary to conduct a simulation experiment. The next section gives brief description about simulation experiment designed in this study.

3. SIMULATION EXPERIMENT

The tool used for the simulation study is OPNET 14.0 modeller. OPNET is a network and application based software used for network management and analysis where it provides variety of simulation samples with Graphical User Interface (GUI) along with the considerable amount of documentation and study cases for wired as well as wireless networks. In this research, 2 different scenarios are used and illustrated through OPNET simulation package. In the first scenario, GRP protocol is simulated and in second scenario RTS/CTS mechanism is simulated. The attributes and parameters set for the creation of the simulation environment for Scenario 1 in OPNET shown on the Table 1 in details below.

Table 1. *Global Parameters for Scenario 1*

| Parameter | Attribute |
|---------------------|-----------------------------|
| Protocol | GRP (AODV based) |
| Simulation Duration | 120 (seconds) with 300 seed |
| Simulation Area | 1000 x 1000 meters |
| Mobility | Random Waypoint |
| Performance Metrics | Throughput, Delay, Load |
| No of Nodes | 15 |

The attributes and parameters set for the creation of the simulation environment for Scenario 2 in OPNET simulation package shown on the Table 2 in details below.

Table 2. Global Parameters for Scenario 1

| Parameter | Attribute |
|---------------------|-----------------------------|
| Protocol | TORA |
| Simulation Duration | 120 (seconds) with 300 seed |
| Simulation Area | 1000 x 1000 meters |
| Mobility | Random Waypoint |
| Performance Metrics | Throughput, Delay, Load |
| No of Nodes | 15 |

The Figure 5 illustrates the simulation setup of a single scenario which can be considered as an global scenario since there is no difference between two scenarios in terms of number of participants on the network where comprising of 15 mobile nodes moving at constant speed of 10 meters per seconds. Total of 2 scenarios have been developed, all of them with mobility of 10 m/s. where a simulation time was taken 300 seconds. Simulation area taken is 1000 x 1000 meters. Packet Inter-Arrival Time (sec) is taken exponential (0.3) and packet size (bits) is exponential (2000). The data rates of mobile nodes are 11 Mbps with the default transmitting power of 0.005 Watts. Random way point mobility is selected with constant speed of 10 meter/seconds and with pause time of constant 100 seconds. This pause time is taken after data reaches the destination only.

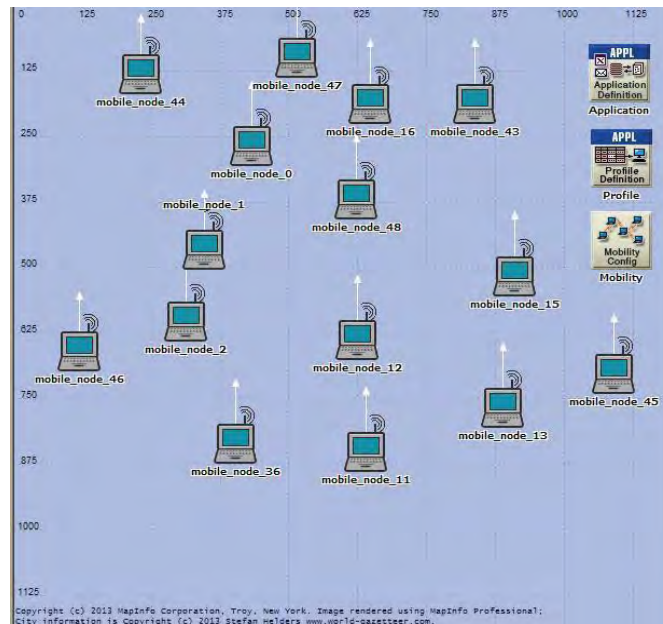


Figure 5. Simulation Environment for 15 nodes

The simulation run time is set as 300 seconds which is equal to 60 minutes with the seed value of 300. Simulation Kernel is set to “optimization”. Application profile, Profile configuration, and Mobility are configured to work on the network in order to provide data flow over the network.

4. SIMULATION RESULTS

The OPNET Simulation package provides two different types of statistics which are known as Global and Object statistics in order to evaluate the performance metrics or the entire simulation outcomes. The global statistics represents the collection of entire network’s data and analysis. The object statistics involves individual nodes statistics. In this simulation experiment, the global discrete event statistics (DES) statistics are taken into consideration for evaluation of performance metrics. This section focuses on results, its analysis and comparison based on the simulation performed in OPNET modeller 14.0.

Based on the parameters set for the simulation scenarios (scenario 1 and scenario 2), the WLAN throughput of both mechanisms are illustrated in the Figure 6.

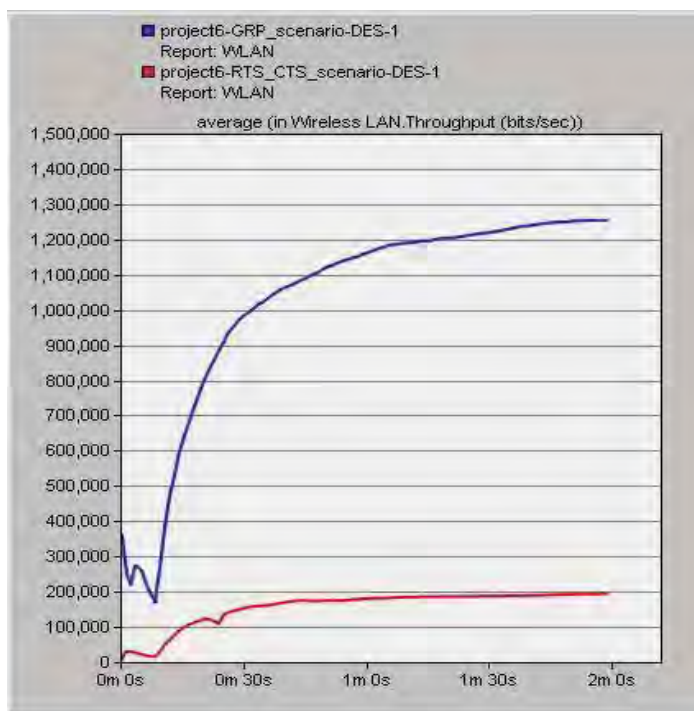


Figure 6. WLAN Throughput of GRP and RTS/CTS Mechanisms

As it is clearly shown on the Figure 6 above, there is significant difference between GRP and RTS/CTS mechanisms. Overall throughput of the network in terms of bits/seconds is significantly decreased when an RTS/CTS mechanism is used with TORA protocol.

Figure 7 below indicates simulation results of average WLAN Delay. the packet end-to-end delay represents the average time in order to traverse the packet inside the network. This includes the time from generating the packet from sender up till the reception of the packet by receiver or destination and expressed in seconds. The Figure 8 below shows the Average WLAN Delay. Since the RTS/CTS mechanism proposes a handshaking mechanism in order to authentication and aims to eliminate hidden node problem and in contrast to this GRP uses CSMA/CA collision avoidance features, the high rate of WLAN delay is expected from RTS/CTS scenario.

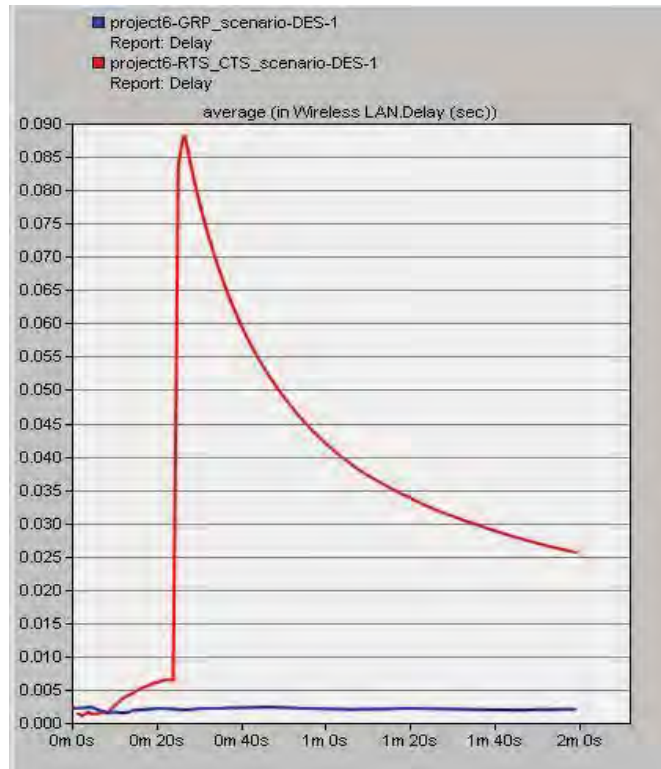


Figure 7. Average WLAN Delay of GRP and RTS/CTS Mechanisms

Finally, the Figure 8 shows the average network load, which is the total traffic received by the entire network from higher layer of MAC that is accepted and queued for transmission. It represents the quantity of traffic in entire network. It indicates the total data traffic in bits per seconds received by the entire network from higher layer accepted and queued for transmission.

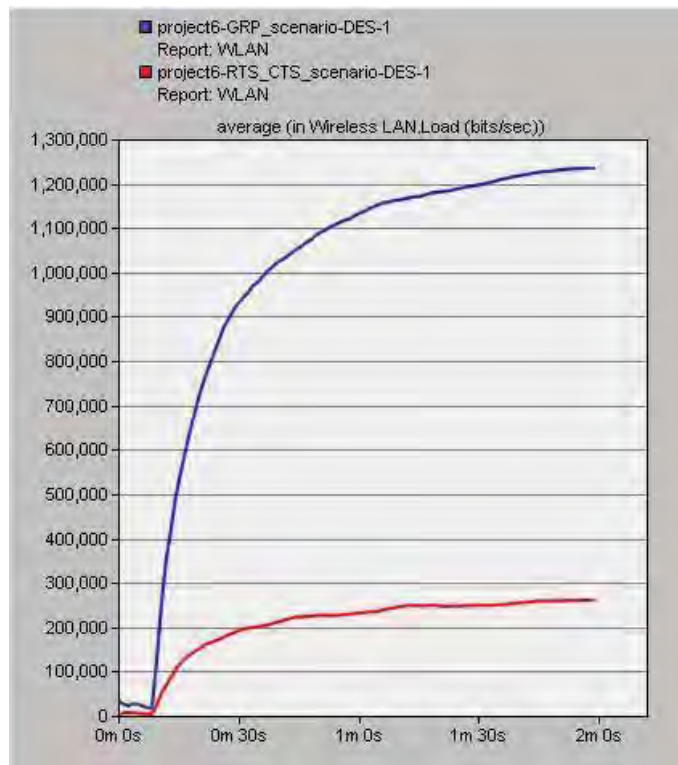


Figure 8. Average network load

As it shown in the Figure 8 above, GRP protocol remains faster and contain much network load than TORA protocol with RTS/CTS mechanism. The high rate of Load represents low rate of data dropped on the network due to collisions or some other reasons. However this may indicate the high rate of overhead problems since there is no differentiation of packages received as broadcasted message.

5. CONCLUSION

The performance of GRP (with simplified MAC protocol) and TORA protocol (with embedded RTS/CTS mechanism) is simulated and compared. The GRP was found to deliver much better performance than the TORA protocol in terms of Throughput, Delay and Network Load. Since high rate of broadcasted packages is delivering to the base station through participating nodes in contrast to TORA protocol, it can be considered that GRP is successful. However, the Base Stations used in GRP protocol are fixed and it is also stated in previous studies. This may lead to expose for a need of a better probability function to work in more flexible situations where mobility is a serious concern for multi-hop wireless networks. As it is mentioned in the previous section, the overhead problem remains as an unsolved issue since there is a constant data flooding mechanism used in GRP with PSC relationship and no differentiation among packages is specified. Researchers should investigate as how to differentiate the forwarding of broadcasted messages to minimize overhead problem for bi-directional data transfers and also in terms of transferred data security.

6. REFERENCES

- [1] Zygmunt Hass, Joseph Y. Halpern, Li Li. "Gossip-based Ad Hoc Routing", In IEEE INFOCOM, June 2002.
- [2] Vamsi K Paruchuri, Arjan Durrezi, Durga S Dash, Raj Jain. "Optimal Flooding Protocol for Routing in Ad-hoc Networks", IEEE Wireless Communications and Networking Conference, March 2003.
- [3] Gruia Calinescu, Ion Mandoiu, Peng-Jun Wan, Alexander Zelikovsky. "Selecting Forwarding Neighbors in Wireless Ad Hoc Networks", Mobile Networks and Application, Volumn9, Issue 2, April 2004.
- [4] Sungho Maeung, Shanchieh Jay Yang, Nirmala Shenoy, "Modelling and Analysis of Gossip Based Relay Networks for Ubiquitous and Extended Access", OPNETWORK 2004, Washington DC 30th August- 3rd September.
- [5] Sari, A., Necat, B. "Impact of RTS Mechanism on TORA and AODV protocol's performance in Mobile Ad Hoc Networks", International Journal of Science and Advanced Technology (IJSAT), ISSN:2221-8386, Vol. 2, No.4, April 2012, pp.188-191.
- [6] OPNET Technologies Inc. "Opnet Simulator," www.opnet.com

Using Data Mining on Linked Data

Yasemin GÜLTEPE

Department of ComputerEngineering, Kastamonu University, Kastamonu, TURKEY

yasemingulpe@kastamonu.edu.tr

ABSTRACT

Semantic web is an extension of the current web that presents semantic data whereby machines and humans can work in collaboration. Linked data is one of the approaches used for acquiring a semantic integrity through bringing together datasets related to a data by creating semantic links between the webpages constituting the semantic web. Linked data allows gathering various data coming from different sources at a single service point. According to the Linked Data principle, the amount of data published in the web environment increases gradually. In the Linked Data network, data mining technology is used in the datasets included in the linked data cloud in order to search for the relations that may allow reaching the information among large-scale data and making predictions about the future. This study aims to put forward recommendations for enhancing the effectiveness of data mining processes on bibliographical sources by using the power of linked data.

Keywords- Linked data, data mining, linked data mining.

1. INTRODUCTION

Data mining is the process of exploring hidden patterns and trends that may allow making predictions about the future by using vast amount of datasets[1]. Main problem in data mining is the application of data mining algorithms within structured databases composed of linked objects. Links between objects are useful for many data mining tasks. However, it is difficult to acquire these links through traditional data mining techniques. In addition, not reliable predictions are provided all the time. Attempts to solve these problems have given birth to a new research area called semantic web.

Semantic web [2], which is regarded as the future of WWW, provides a shared model for reusing and sharing the information. The semantic web is a web environment that enables well-defined information and services to be easily understandable by machines. In addition, the semantic web is an extension of the current web that presents semantic data whereby machines and humans can work in collaboration.

The semantic web is based on data and linked data on the basis of relations between data. At the present time, linked data technology makes it easier and faster to collect vast amount of data from different data sources and to produce and merge data for particular purposes. Linked data [3, 4] is one of the approaches used for acquiring a semantic integrity through bringing together datasets related to a data by creating semantic links between the webpages constituting the semantic web. Web of Data represents all linked data. While the current World Wide Web (WWW) provides a web that can be understood by humans alone, data network aims to create a web that can be read and understood by machines

According to the Linked Data principle, the amount of data published in the web environment increases gradually. Linked data allows gathering various data coming from different sources at a single service point. In the Linked Data network, data mining technology [1, 5, 6] is used in the datasets included in the linked data network in order to search for the relations that may allow reaching the information among large-scale data and making predictions about the future.

Bibliographic sources are the secondary sources that do not make us reach the information directly, but enable us to be aware of the existence of different kinds of information sources that contain the information we look for. They are the directional sources that give descriptive information/bibliographic identity concerning the information sources they contain and guide users through reaching main sources. This study aimed to put forward recommendations for enhancing

the effectiveness of data mining processes on bibliographic sources. These data are of big importance for the preparation of scientific research and bibliography in particular.

This study focused on the importance of data mining technology to reach the information over the semantically linked data in the linked data network and introduced the architecture of linked data mining. Thus, the second section of the paper deals with the related previous studies. The third section introduces the architecture of linked data mining. The fourth section summarizes the results.

2. RELATED WORKS

Obtaining and analyzing data through data mining techniques has become important in academic, commercial, sociological, etc. fields. Linked data can be used in data mining as an additional information source that supports the interpretation of the information obtained via data analysis. There are some approaches that make use of linked data technologies in order to interpret data mining results [1, 5, 6, 7]. In general, these approaches facilitate the interpretation of information by selecting data or creating sub-data examples to be discovered for the analysts who are to conduct data mining.

Today, availability of opportunities to use linked data and data mining technologies in many application areas and for various purposes allows integrating the data collected from many different sources through different methods. One of these application areas is health. Some studies have been conducted to increase the efficiency of such processes as finding and linking the suitable data, drug monitoring, and managing drug interaction[8].

At the present time, web basically consists of the linkage of web documents based on HTML documents to one another. Linked data, on the other hand, keeps the data in documents as RDF¹ (Resource Description Framework) and links data (i.e. concepts) instead of documents [4]. RDF is a data model that provides field-independent formal semantics in regard to graph sources.

Some studies [6, 9] have used the semantic web query language called SPARQL in order to query linked data network and monitor the variations between RDF links. In addition, the number of queries can be reduced by creating easy-to-use SPARQL query templates. There are various ways to make data meaningful and to reveal the hidden meanings within data. These studies are basically about data discovery on linked data network. Nebotand Berlanga (2010) suggested a method that could be used in the association rules analysis—one of the first techniques used on RDF-based data in data mining. This method employed ontology axioms and semantic tags in order to produce semantic processes out of semantic web data.

Open source FeGeLOD can be used for data mining applications on the data in the linked data network [10]. FeGeLOD is based on unsupervised learning method. Specific functions are applied through creating natural sets out of input data to infer the basic properties of the datasets in the linked data network.

Link mining is an example of multi-relational data mining. It involves various tasks including but not limited to predictive and descriptive modeling. Among such tasks are link-based classification, link-based set interpretation, definition of link types, prediction of link power, the number of links, and register links. Link mining tasks are used in applying data mining to the datasets of social networks emerging over the Internet. Graphic display is frequently employed in the display of social networks. Link mining is an example of data mining in multi-relational datasets illustrated in one graph[11].

3. DATA MINING ON LINKED DATA

Linked data introduces an approach based on web standards for data integration. The Linked Open Data (LOD) project [4] was launched. It is open to anyone who wants to publish data and make semantic inferences out of data according to linked data principles. Linked data in different application areas enlarge continuously and become linked open data cloud.

¹<http://www.w3.org/RDF/>

This study employed bibliographic data sources as linked open data cloud. Bibliographic sources are the secondary sources that do not make us reach the information directly, but enable us to be aware of the existence of different kinds of information sources that contain the information we look for. Bibliography involves papers, books, authors, institutions, articles, conferences, etc.

This study aimed to put forward recommendations for enhancing the effectiveness of data mining processes on bibliographic sources by using the power of linked data. Various dictionaries (Dublin Core, FOAF vs.) and ontologies ontolojiler (BIBO, DBLP vs.) can be used for presenting bibliographic sources. Data mining refers to the use of algorithm for exporting the models and information produced by Knowledge Discovery in Databases (KDD) processes.

Narasimhaand Vyas (2011) made some changes in KDD for the data mining applications that could be used within the framework of the requirements of semantically linked data on the linked data cloud. As is seen in figure 1, data mining stages were defined on the linked data [5]. Finding people conducting joint academic studies or discovering research areas and joint publication areas can be given as examples of the application of data mining in the bibliographic field.

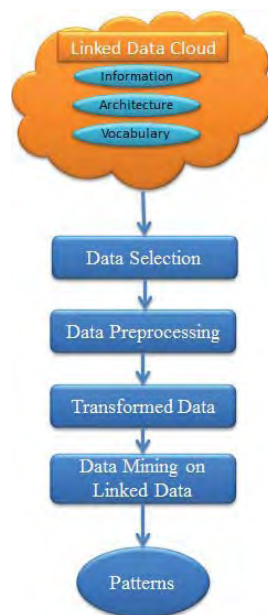


Figure 1. Data Mining on Linked Data.

Data mining process and stages for linked data can be explained as follows [5]:

3.1. Data Selection

The concepts HTTP and URI underlie linked data. Linked data keeps the data in documents as RDF and links the data (i.e. concepts) rather than documents. Data sources are named with the reference URI. When a linked data is published, a good URI must be selected for sources. That is important for other data providers to establish a sound link for other data sources. The data source can be accessed via HTTP. The linked data technology prescribes data to be able to be queried via the query language SPARQL through the protocol HTTP [12]. Figure 2 gives a sample query defined for obtaining the publications containing the word “semantic” in their titles as well as the authors of such publications.

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX bibo: <http://purl.org/ontology/bibo/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX doco: <http://purl.org/spar/doco/>
  
```

```

SELECT ?article ?author
WHERE {
  ?article rdf:type bibo:AcademicArticle .
  ?article dc:isPartOf ?journal .
  ?article bibo:author ?author .
  
```

```

    ?authorfoaf:name      ?name .
  ?articledoco:contains?section .
  ?sectiondc:title       ?title .
  FILTER (regex(str(?title), "semantic")) . }

```

Figure 2. SPARQL Query Example.

3.2. Data Preprocessing

Since the quality of data affects results, it is quite important that the data to be used are subjected to preprocessing. Quality data produce quality outputs. Data preprocessing plays an important role in data mining [13]. Many data preprocessing techniques are available. They are data merge, data filtering, and data clustering/partitioning.

3.2.1. Data Integration

In the first step of linked data mining, data are obtained from the data sources existing in the data cloud. Distributed data in different geographical positions need to be merged coherently. Data are merged on the basis of some common relations indicated in data sources [14]. A user may want to work with different data sources. In this case, data sources can be selected based on different factors. For instance, a study is conducted to list the Twitter accounts of those researchers who work in the field of “Semantic Web”. Firstly, the query result obtained in Section 3.1 is used. Then Twitter data source is queried for obtaining information about the Twitter accounts of academic staff. The obtained information is merged in order to find the answer of the related question.

3.2.2. Data Filtering

When data from different data sources are merged, abundant data may emerge. For example, the obtained list of authors contains “Y. Y. Gultepe” and “Yasemin Y. Gultepe” as two author names. However, both of them are the name of the same author. In this step, such author names are eliminated.

3.2.3. Data Clustering / Partitioning

Among the clustered or partitioned data, those which are similar by nature are grouped together. By this means, it is aimed to predict the datasets to which each data belongs to and thus ascertain the data displaying abnormal behaviors. For example, the population data of a country can be divided into certain sub-classes: low population density cities, medium population density cities, and high population density cities.

3.3. Transformed Data

In this step, initial data are transformed into final data that are to provide a basis for works. In the linked data, some quality types may not be suitable for the data mining algorithms to be applied. To make them suitable, source data may be converted into different formats or values.

3.4. Data Mining on Linked Data

Results are obtained by applying data mining methods to such processed data. Data methods can basically be grouped as classification, clustering and association rules. The data obtained through the application of data mining methods need to be presented to decision-makers who are to evaluate them.

4. RESULTS

Linked data is based on the publication of data over the web in the form of datasets compatible with the RDF standard and the association of different datasets with RDF links. There is a general consensus among researchers that the use of linked data for data mining may be rewarding. Studies concerning the use of linked data for those data which are kept in data sources, but have not been discovered yet are generally conducted within the scope of classic KDD.

The use of linked data in data mining processes may shorten the data preparation (the most demanding part of data mining) time and decrease costs through accelerating the interpretation of the results after such preparation period. The fact that linked data semantically link the concepts within datasets with one another makes the use of linked data in data mining processes easier.

Narasimha and Vyas suggested a system involving an interaction of linked data and data mining techniques that converted data which were large in amount and existed in different formats and locations (heterogeneous and distributed) into information and made them meaningful. In the present study, this system was used in the field of bibliographic data. This system makes important contributions to the solutions of the problems encountered in the mining of bibliographic records.

5. REFERENCES

- [1] E.L. Mencia, S. Holthausen, A. Schulz, F. Janssen, (2013), Using Data Mining on Linked Open Data for Analyzing E-Procurement Information, Data Mining on Linked Data workshop at ECML/PKDD2013.
- [2] T. Berners-Lee, J. Hendler, O. Lassila, (2001), "The Semantic Web", Scientific American, 284(5), pp. 34-43.
- [3] T. Heath, C. Bizer, (2011) "Linked Data: Evolving the Web into a Global Data Space", Synthesis Lectures on the Semantic Web: Theory and Technology, Morgan & Claypool Publishers, 136p.
- [4] C. Bizer, T. Heath, T. Berners-Lee, M. Hepp, (2009), "Linked Data-The Story So Far", International Journal on Semantic Web and Information Systems, 5(3), pp.1-22.
- [5] R.V. Narasimha, O.P. Vyas, (2011), "LiDDM: A Data Mining System for Linked Data. In: workshop on Linked Data on the Web", Proceedings of the LDOW2011.
- [6] E.L. Mencia, S. Holthausen, A. Schulz, F. Janssen, (2013), Using Data Mining on Linked Open Data for Analyzing E-Procurement Information, Data Mining on Linked Data workshop at ECML/PKDD2013.
- [7] Z. Akhtar, M.K. Singh, N. Begam, (2012) "Challenges and Usage of Link Mining to Semantic Web", International Journal of Electronics and Computer Science Engineering, 8(1), pp. 775-780.
- [8] J. Lorey, F. Naumann, (2013), "Detecting SPARQL Query Templates for Data Prefetching", ESWC, volume 7882 of Lecture Notes in Computer Science, pp. 124-139.
- [9] Z. Akhtar, M.K. Singh, N. Begam, (2012) "Challenges and Usage of Link Mining to Semantic Web", International Journal of Electronics and Computer Science Engineering, 8(1), pp. 775-780.
- [10] H. Paulheim, J. Furnkranz, (2012), "Unsupervised generation of data mining features from linked open data", International Conference on Web Intelligence and Semantics.
- [11] Z. Afshan, M.S. Sarwar, (2014), "Link Mining in Social Networks", IJSRD, 1:12, pp. 2826-2831.
- [12] A. Nolle, G. Nemirovski, Á. Sicilia, J. Plequzelos, (2013), An Approach for Accessing Linked Open Data for Data Mining Purposes, Proceedings of Ramid Miner Community Meeting and Conference.
- [13] H. Paulheim, (2013), "Exploiting Linked Open Data as Background Knowledge in Data Mining", In Workshop on Data Mining on Linked Open Data, pp.1-10.
- [14] W.P. Norman, K. Christodoulou, A.A.A. Fernandes, B. Parsia, C. Hedeler, (2012), "Pay-as-you-go data integration for linked data: opportunities, challenges and architectures", SWIM 2012.

Kidney Segmentation From Abdominal CT Images By Using Connected Component Labeling Algorithm

Seda Arslan TUNCER¹, Ahmet ALKAN²

¹ Department of Informatics, Firat University, Elazig, Turkey

² Department of Electrical & Electronics Engineering, KSU, Kahramanmaras, Turkey

¹satuncer@firat.edu.tr, ²aalkan@ksu.edu.tr

ABSTRACT

Abdominal Computer Tomography (CT) scan data are used in many clinical studies as a correct and efficacious help for diagnosis and treatment. Segmentation of viscera on abdominal imaging facilitates diagnosis and focusing on the areas of interest. Among abdominal imaging tasks, kidney segmentation is a difficult one given the proximity of different organs and the similarities between abdominal tissues. This study proposes a fully automated approaches to kidney segmentation and discusses the degree of success of them . The proposed methods essentially consist of three stages. Since the spine was used as reference in the study, the images were first treated to define the coordinates of the spine. The application of filters to the images provided a differentiation of closely connected intrarenal cortical and medullary areas and the clear delineation of the spine. In the second stage, kidney fields were obtained using the Connected Component Labeling (CCL). The performance evaluation of the used algorithms was done based on the algorithm-based and the manual segmentation results. The criteria used to confirm the accuracy of the segmentation operation was those of the Zijdenbos similarity index (ZSI). ZSI is an assessment index based on the difference between the area segmented by the segmentation algorithms and the area segmented by a specialized physician, both represented in pixels. These assessments resulted in a success rate of 85.2 % for right kidney and 86.8 % for left kidney, respectively, by the CCL algorithms. This system is expected to be helpful both in clinical diagnosis and in medical training. Results indicate that the proposed segmentation procedure may be used to assist medical diagnosis and decision.

Keywords – Kidney segmentation, CCL, Abdominal images, Image processing, ZSI

1. GİRİŞ

Tıbbi görüntüleme ile anatomi hakkında detaylı bilgiler elde edilebildiğinden, tanı amaçlı görüntüleme birçok açıdan önem arz etmektedir. Görüntüleme cihazları tarafından sunulan veriler, görüntü işleme yöntemleriyle tüm veri yerine ilgilenilen doku alanına ulaşımı sağlayarak kolay işlenebilir hale gelmektedirler. Bölütleme ve görüntüleme birbirleriyle yakın ilişkili iki alan olup, birçok radyolojik uygulamada kullanılmaktadırlar. Hem bölütleme, hem de üç boyutlu görüntüleme için, yeni tekniklerin geliştirilmesi kadar, Bilinen yöntemlerin uygulanabilir hale getirilmeside önemlidir. Radyolojide kullanılan görüntüleme cihazlarından elde edilen verilerde yetersiz çözünürlük gibi problemlerden dolayı bazı zamanlarda hekimler tanı koymakta zorlanmaktadırlar. Bu nedenle görüntülerin işlenip daha anlamlı hale getirilmesi gerekmektedir. Farklı fazlardan elde edilmiş, kontrastı artırılmış Bilgisayarlı Tomografi (BT) görüntülerinin bölütleme yöntemi böbrekteki lezyonların bulunmasını ve karakterize edilmesini sağlayan etkili bir yöntemdir. Böbrek görüntüsü hastanın nefesini tutmasından önce ve hastaya kontrast madde enjekte edildikten sonra alınmaktadır. Batın bölgesi burada tekrar tekrar kontrast maddenin aktarımı sırasında taranır. Kontrast madde aktarımı sırasında alınan bu görüntülerin farkı alınarak lezyon tespiti yapılmaktadır. Bölütleme, görüntüdeki herhangi bir örüntüyü veya görüntünün herhangi bir parçasının geri kalan parçalardan ayrılması işlemine denir [1]. Medikal görüntü işleme alanında bölütleme, anatomik yapıların çalışılmasında, tedavi planlamasında, bilgisayar destekli ameliyatlarda sıklıkla kullanılmaktadır [2]. Böbrek bölütleme, böbreğin korteks ve meulla

alanlarının farklı yoğunluğa sahip olması, bazı görüntülerde karaciğere yapışık ve gri seviyelerinin benzer olması sebebiyle zorlaşmaktadır. Geliştirilen bazı yöntemler yarı otomatik kullanıcı etkileşimli yöntemler olup bu yöntemlerin dezavantajı kullanıcıya zaman kaybettirmesidir. Bu nedenle otomatik olarak bölütleme yapmak her zaman uzmanlarca tercih edilmektedir.

Abdominal görüntülerden organ bölütlemeye ilgili literatürde birçok çalışma bulunmaktadır. Bilgisayarlı tomografik planlama sağ ve sol böbreklerin bölütlenmesi için geleneksel kullanıcı güdümlü bir yapı önerilmiştir. Deforme olabilen şekil modelleri bölütleme için kullanılmıştır [3]. Başka bir çalışmada tıbbi görüntülerden üç boyutlu anatomik hacimleri bölütleme için deforme model tabanlı bir yaklaşım sunulmuştur. Çalışmada bölütleme ile birlikte modelleme işlemi de yapılmıştır [4]. Diğer bir çalışma da gri seviye böbrek görüntüsünden istatistiksel bilgiler ile temsil edilmiş ve bu bilgilerle birlikte deforme olabilen bir modeller kullanılmıştır [5]. Böbrek tümörleri üzerine yapılan bir çalışma da örnek görüntüler üzerinde yapılan doku analizine dayanarak, böbrek tümörü bölge büyütme algoritmasının başlangıç noktası olarak kullanılmıştır. Çalışmada abdominal BT görüntüleri sayısallaştırılmış ve gri-seviyeli eşikleme yöntemi böbrek bölütleme için kullanılmıştır [6]. BBE (Bağlantılı Bileşen Etiketleme) ve bölge büyütme yönteminin birlikte kullanıldığı başka bir çalışmada ilk olarak uyarlanı eşikleme kullanılmıştır. Daha sonra BBE kullanılarak böbrek bölütlenmesi için tahmini bir başlangıç noktası elde edilmiştir. Elde edilen başlangıç noktaları kullanılarak bölge büyütme yöntemiyle bölütleme işlemi yapılmıştır [7]. Böbreğin otomatik olarak bölütlenmesi ve bölütlenmiş farklı fazlardaki böbreklerin çakıştırılmasını sunulduğu başka bir çalışmada böbreğin otomatik bölütlenmesi, hesaplanan omurga konumuna göre morfolojik operatörlerle yapılmıştır. Omurga konumu Daw-Tung Lin ve diğerlerinin [8] bulduğu matematiksel formulasyona göre hesaplanmıştır. Çakıştırma işlemi için matlab kütüphanesinden farklı yöntemler denenmiştir. Dikkate değer bu çalışmalar haricinde böbrek bölütleme ile ilgili çalışmalar devam etmektedir. Yapılan tüm çalışmalardaki amaç böbrek bölütleme işleminin yüksek başarı ile elde edilmesidir. [9,10,11]

Bu çalışmada otomatik olarak bölütlemeyi gerçekleştirebilmek için batın bölgesinin anatomik özelliğinden faydalanılmıştır. Omurga bölütlendikten sonra omurganın koordinatlarından faydalanılmış sağ ve sol böbrek alanları ayırılmıştır.

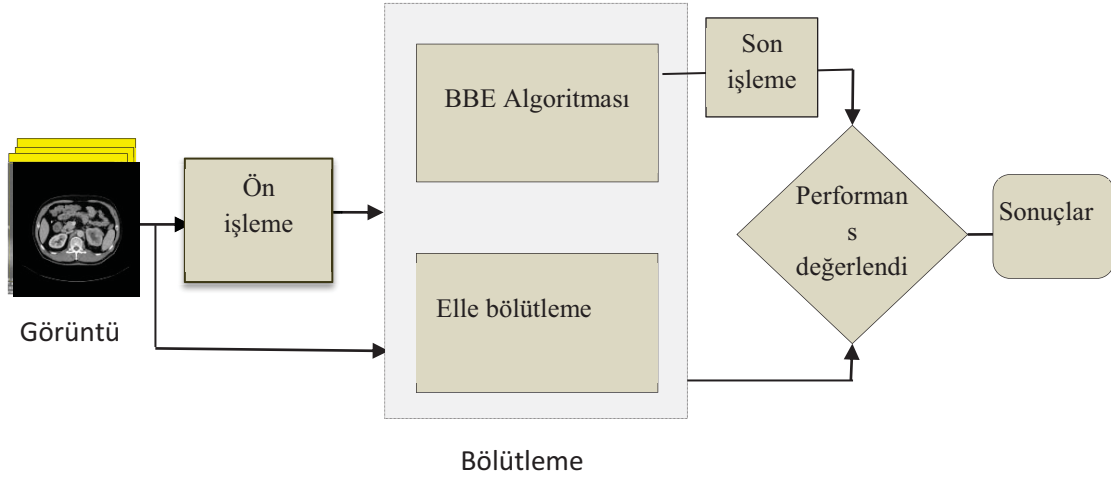
2. MATERYAL VE METOT

2.1 Materyal

Fırat Üniversitesi Tıp Fakültesi radyodiagnostik anabilim dalı görüntü arşivleme sisteminden 50 hastaya ait görüntü serisi alınmıştır. İlgilenilen alan tüm seride bulunmadığından serilerdeki tüm görüntüler üzerinde çalışılmamıştır. Görüntüler, damardan opak madde verilmesini takiben portal fazda çekilmiş BT görüntüler olup 5 mm kesit kalınlığı ile çekilmiştir. Görüntüler DICOM formatındadır. Elde edilen görüntülerde yer alan böbrek görüntüleri uzman bir radyolog tarafından manuel olarak bölütlenmiş ve referans görüntüler oluşturulmuştur.

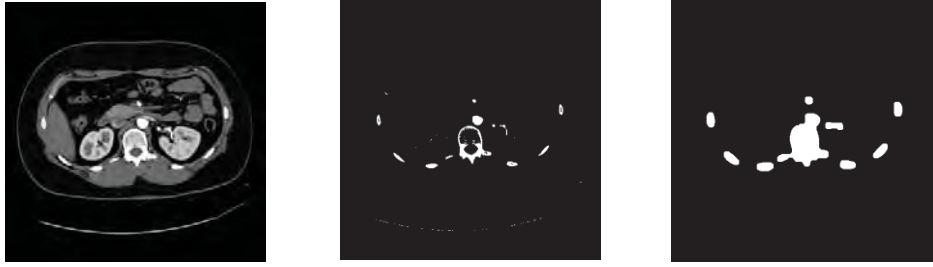
2.2 Metot

Bu çalışmada abdominal görüntülerden böbreklerin BBE yöntemi kullanılarak bölütlenmesi hedeflenmiştir. Çalışmanın amacı abdominal görüntülerden böbreği bölütleyerek böbrek üzerinde kanserli hücre ya da lezyon belirleme çalışmalarına zemin hazırlamaktır. Çalışma temel olarak 3 aşamadan oluşmaktadır. Ön işleme aşamasında matematiksel morfoloji kullanılmıştır. Morfoloji temel küme işlemlerine dayanan, görüntüdeki sınırlar, iskelet gibi yapıların tanımlanması ve çıkartılması, gürültü giderimi, bölütleme gibi uygulamalar için gerekli bir araçtır. Görüntü işlemede genellikle, morfolojik süzgeçleme, inceltme, budama gibi ön veya son işlem olarak kullanılırlar. Daha sonra morfolojik işlemlerden geçen veriler BBE yöntemi kullanılarak bölütlenmiştir. Bölütleme sonrasında veriler son işlemden geçirilerek ilgilenilen alan elde edilmiştir. Çalışmanın aşamaları Şekil 1 de verilmektedir. [12].



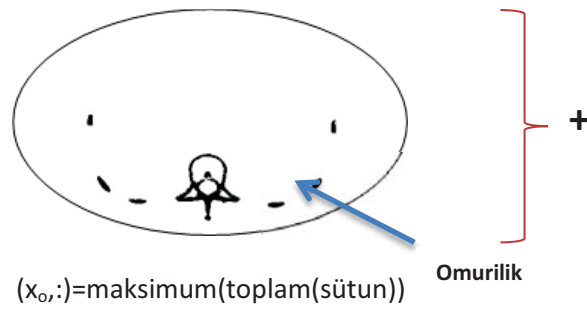
Şekil 1. Yapılan çalışmanın aşamaları

Çalışmanın ön işleme sonrasında omurga elde edilmiştir. Böylece sağ ve sol böbrek için ayrılma noktası belirlenecektir. Şekil 2 omurganın adım adım elde edilmesini göstermektedir.



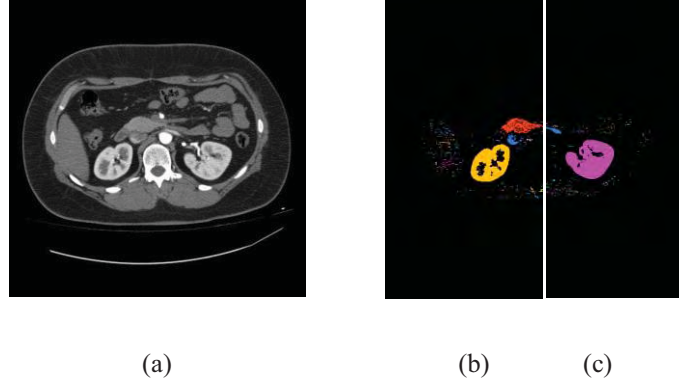
Şekil 2. Omurganın adım adım elde edilmesi

Çalışmada omurga morfolojik işlemlerle tespit edildikten sonra, sütun toplamları hesaplanmıştır. Toplamı maksimum olan sütunun omurga orta noktası olduğu kabul edilmiş ve görüntü bu noktadan ikiye ayrılmıştır. Bu nokta ile sağ ve sol böbreğin ayrı ayrı bölütleneceği alanlar hedeflenmiştir. Şekil 3 Sağ ve sol böbrek için ayrılma noktasının belirlenmesini göstermektedir.



Şekil 3. Sağ ve sol böbrek için ayrılma noktasının belirlenmesi

Böbrek bölütleme işlemi için ikinci aşama BBE yönteminin kullanılmasıdır. BBE, görüntüde birleşik olan nesnelere etiketleyip komşu olan pikselleri bir grup içerisinde toplayarak birbirinden ayırt edebilecek hale getirir. Gruplama sonrasında, resim üzerindeki her bir grup bir nesneyi temsil edecek şekilde numaralandırılır. Böylece istenen grup numaralı nesneyi diğer nesnelere ayırmak suretiyle işlem yapılabilir. BBE yöntemi 4-komşuluk ve 8-komşuluk olarak ikiye ayrılır. Bir çok uygulamada 8-komşuluk tercih edildiğinden kodlarda çapraz pikseller de komşu olarak işleme dahil edilir [13], [14], [15]. BBE algoritması sonrasında her bir ayrı parça için R,G,B değerleri elde edilir. Belirli kanallarda en büyük bileşenin böbrek olmasından yararlanılarak, bölütlenmiş böbrek görüntülerine ulaşılır. Elde edilen sonuçlar 1. aşamada elde edilen bölünme noktasından tekrar birleştirilerek 512x512 lik orijinal boyutlarda görüntü elde edilir. Şekil 4. a) orijinal görüntüyü 4.b) BBE sonunda elde edilmiş R kanalını 4.c) BBE sonunda elde edilmiş B kanalını göstermektedir. Elde edilen bu görüntülere son işlem uygulandıktan sonra bölütleme işlemi tamamlanmıştır.



Şekil 4. R ve B kanalındaki görüntüler. (a) Orijinal görüntü , (b) R kanalındaki görüntü, (c) B kanalındaki görüntü

4. PERFORMANS DEĞERLENDİRME

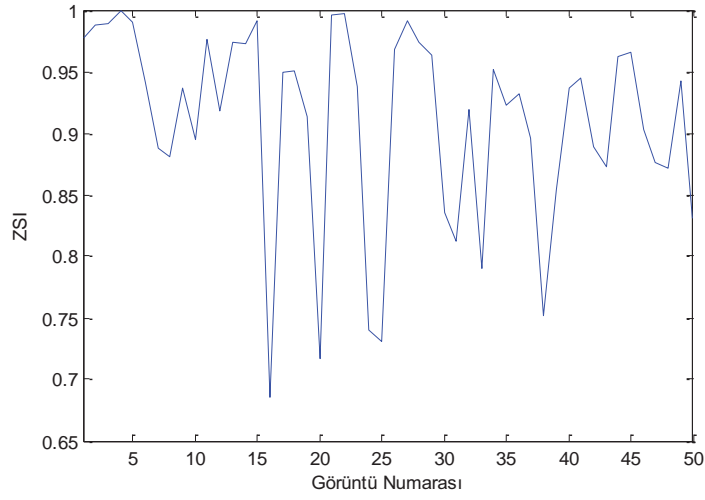
Bölütleme işleminin başarısını değerlendirmek için Zijdenbos Similary İndex kriterleri kullanılmıştır. Çalışmanın performansı 50 hastadan alınan Bilgisayarlı tomografi görüntülerindeki sağ ve sol böbrek için ayrı ayrı elde edilmiştir. Çünkü sağ ve sol böbreği çevreleyen organlar farklılık göstermektedirler.

4.1 Zijdenbos Similary İndex

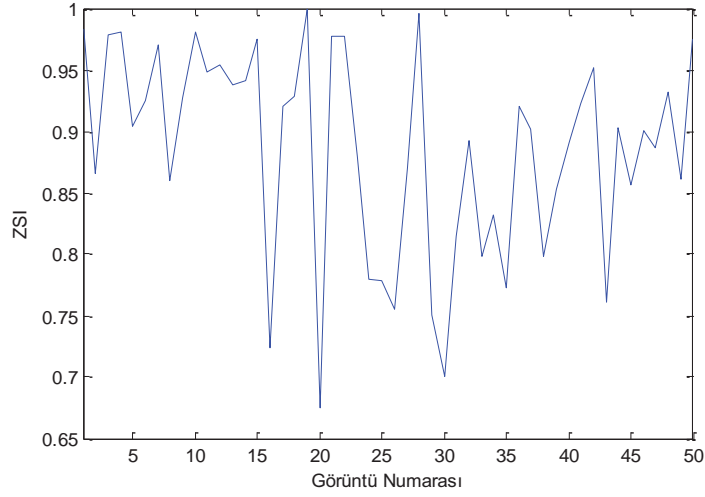
ZSI (Zijdenbos Similary İndex) bölütleme algoritmaları ile pixel cinsinden bölütlenen alan ile uzman hekim tarafından pixel cinsinden bölütlenen alan arasındaki farka dayalı değerlendirme tekniğidir. ZSI değeri 0 ve 1 arasındadır. İndeks değerinin 1'e yakınlığı uygulamada sonucun başarılı olduğunu ifade eder. ZSI indeksinin 0.7 den büyük olması mükemmel sonuç olarak kabul edilir [16,17]. A ve M yukarıda tanımlanan verilen pixel cinsinden alan değerlerini göstermek üzere ZSI indeksi denklem 1'deki gibi verilir.

$$ZSI = 2 * \frac{(A \cap M)}{(|A| + |M|)} \quad (1)$$

Şekil 5 ve şekil 6 sırasıyla sağ ve sol böbrek için BBE algoritması ile elde edilen bölütleme sonuçlarının ZSI değişimini göstermektedir.



Şekil 5. Sağ böbrek için elde edilen ZSI sonuçları



Şekil 6. Sol böbrek için elde edilen ZSI sonuçları

5. SONUÇLAR

Bu makalede kontrastı artırılmış Bilgisayarlı tomografi görüntülerinde böbrek bölütlemeye için BBE algoritması kullanılmıştır. Makalede kullanılan yöntemin doğruluğunu göstermek amacıyla Zijdenbos Similarity İndeksi kullanılarak performans değerlendirilmesi yapılmıştır. Hem manuel hem de önerilen yöntemlerle elde edilen bölütleme sonuçları karşılaştırıldığında BBE algoritması ile böbrek bölütlemeye kabul edilebilir sonuçlar elde edilmiştir. Sonuçlar, önerilen bölütleme yapısının hekimin hasta ve hastalık hakkında tanı ve karar verebilmesi aşamalarında kullanılabileceğini göstermiştir. Çalışmanın bundan sonraki aşaması böbreklerdeki lezyonları tespit etmek olacak ve yapılan çalışma bundan sonraki çalışmaların ilk aşamasını oluşturacaktır

6. REFERANSLAR

- [1] R. C. Gonzales, R. E. Woods, Digital Image Processing, PrenticeHall, 2001.
- [2] Dzung L. Pham, Chenyang Xu, and Jerry L. Prince. "Current methods in medical image segmentation", Vol. 2: 315-337 (Volume publication date August 2000)

- [3] J. Z. Chen, G. S. Tracton, M. Rao, S. Joshi, E. L. Chaney, and S. M. Pizer, "Comparison of automatic and human segmentation of kidneys from CT images," *Int. J. Radiation Oncology Biology Physics*, vol. 54, p. 82, 2002.
- [4] B. Tsagaan, A. Shimizu, H. Kobatake, K. Miyakawa, and Y. Hanzawa, "Segmentation of kidney by using a deformable model," in *Int. Conf. Image Processing*, vol. 3, Thessaloniki, Greece, 2001, pp. 1059–1062.
- [5] B. Tsagaan, A. Shimizu, H. Kobatake, and K. Miyakawa, "An automated segmentation method of kidney using statistical information," in *Proc. Medical Image Computing and Computer Assisted Intervention*, vol. 1, 2002, pp. 556–563.
- [6] Daw-Tung Lin, Member IEEE, Chung-Chih Lei, and Siu-Wan Lung, "Computer-Aided Kidney Segmentation on Abdominal CT Images" *IEEE Transaction on Information Technology in Bio Medicine*, Vol. 10, No. 1, January 2006.
- [7] Gao Yan; Boliang Wang, "An automatic kidney segmentation from abdominal CT images," *Intelligent Computing and Intelligent Systems (ICIS)*, 2010 IEEE International Conference on, vol. 1, no., pp. 280, 284, 29-31 Oct. 2010
- [8] Akyar H., Selver M. A., Demir K. G. "Kontrastı Arttırılmış Batın MR Görüntülerinden Böbrek Bölütlenmesi ve Çakıştırılması", *SIU2008* no:330
- [9] Benali H, Buvat I, Frouin F, Bazin JP, DiPaola R. A statistical model for the determination of the optimal metric in Factor Analysis of Medical Image Sequences (FAMIS). *Phys. Med. Biol.* (1993) 38 : 1065-1080
- [10] Buvat I, Benali H, Frouin F, Bazin JP, DiPaola R. Target apex-seeking in factor analysis of medical image sequences. *Phys. Med. Biol.* (1993) 38 : 123-138
- [11] Tsagaan B, Shimizu A, Kobatake H, Miyakawa K. An Automated Segmentation Method of Kidney Using Statistical Information. *MICCAI (2002) Lecture Notes in Computer Science* 2488 : 556–563
- [12] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, —*Digital Image Processing Using MATLAB*
- [13] Min Li; Xiaolin Zheng; Xiaoping Wan; Hongyan Luo; Shaoxiang Zhang; Liwen Tan, "Segmentation of brain tissue based on connected component labeling and mathematic morphology," *Biomedical Engineering and Informatics (BMEI)*, 2011 4th International Conference on, vol. 1, no., pp. 482, 485, 15-17 Oct. 2011 doi: 10.1109/BMEI.2011.6098294
- [14] Michael B. Dillen court and Hannan Samet and Markku Tamminen (1992). "A general approach to connected-component labeling for arbitrary image representations". *J. ACM.*
- [15] <http://www.cse.unr.edu/~bebis/CS791E/Notes/ConnectedComponents.pdf>
- [16] Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of White matter lesions in MR images: Method and validation. *IEEE Trans Med Imag.* 1994 Dec;13:716–24
- [17] Alkan A., Arslan Tuncer S, Gunay M., "Comparative MR image analysis for thyroid nodule detection and quantification, *Measurement*, Volume 47, Pages 861-868, January 2014

ARTIFICIAL NEURAL NETWORK BASED POWER DISTRIBUTION SYSTEM MODELLING AND HARMONIC ESTIMATION

Suat ÖZDEMİR¹ , Metin DEMİRTAŞ² , Serkan AYDIN³

¹Celal Bayar Üniversitesi, Turgutlu MYO, Elektrik Bölümü,

Turgutlu/Manisa, TÜRKİYE

²Balıkesir Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Müh. Bölümü,

Balıkesir, TÜRKİYE

³ Celal Bayar Üniversitesi, Teknoloji Fakültesi, Mekatronik Müh. Bölümü,

Turgutlu/Manisa, TÜRKİYE

¹suat.ozdemir@cbu.edu.tr, ²mdtas@balikesir.edu.tr, ³serkan.aydin@cbu.edu.tr

ABSTRACT

This study presents an Artificial Neural Networks (ANN) based modeling and harmonic estimation system of power distribution system. The importance of power quality in Electrical systems increases with each passing day and has become an important criterion in energy system. The harmonics causes resonance, sudden trips, loss and these problems adversely affect the quality of power. In this study THD_I values of residential and industrial zones were examined. In order to get an estimate of the value of THD_I values ANN were used. With this method, accurate estimated values were obtained from the ANN model. With the proposed ANN model, equipment that improves the power quality can be preselected while planning energy distribution systems. In this way, by reducing harmonic values, a significant improvement in power quality can be achieved. This model can be used for the planning the energy systems which has the high Power Quality and to estimate THD_I values in different systems.

Keywords : *Power Distribution System, Artificial Neural Network, Harmonic , Power Quality, THD*

1. INTRODUCTION

With the increasing the technology the nonlinear loads are applied in Power distribution system widely. Devices that using semiconductor switching elements caused to increase grid connected non-linear load. Household devices like television, computer, air conditioning; Industrial type load like UPS, Speed control device, welding machine , can be given as an example to this type of non-linear load. This type devices that generates harmonic, negatively affects the power quality. Harmonics make adverse effects in transmission lines, loads and electrical equipment because of this effects harmonics cause power quality problems [1]. Harmonics causes the power quality problems such as losses in transformers, trenches, Disruption in communication systems, sudden release, losses in power distribution [2], [3]. Total Harmonic Distortion (THD) is used as an indication of current and voltage distortion values. Harmonic limitations are described in the IEEE 519 standard [4], [5]. In addition, reference [6] and [7] describes the harmonic distortion limit values in IEC standards. ANN's, are frequently used for harmonic estimates and signal analysis [8], [9]. ANN's are used In Power systems Transients, [10], and obtaining the interharmonics [11]. In Reference [12] the harmonics amplitude and the angle estimation has been done with using the ANN estimation method. The harmonic estimation were made with very high accuracy [13]. In this study, THD_I values in various regions were detected by measurement devices connected to the grid. These values have been used by ANN for training and testing purposes. In the second part of the study, general information about harmonics, in the third part harmonic measurements, In the fourth section simulation and the results are presented in the last section. In reference [14] Artificial neural networks have been applied to

estimate the load. In reference [15] the ANN were used to estimate the nonlinear harmonic loads. In reference [16] the dynamic system harmonic analysis is performed using ANN. In reference [17] the ANN were used in the analysis of harmonic distortion..In Reference [18] ANN's estimation method has been proposed for the current estimate of nonlinear loads which contains harmonic components. With the purpose of quickly harmonic determinations the ANN systems were used [19].

2. POWER SYSTEM HARMONICS AND MEASUREMENTS

2.1 Power systems harmonics

Harmonics causes the distortion in the energy system. The increase in nonlinear loads and circuit elements, could cause to unstable and low quality grids. Harmonic voltages and currents in an electric power system result of non-linear electric loads. Because of nonlinear loads in the system, sinusoidal waves can occur at different frequencies. Harmonics and harmonic analysis are needed the determination to reduce the negative effects of nonlinear loads to power systems. If there are Nonlinear loads on systems it is necessary to redefine the electrical parameters.

Total Harmonic Distortion THD is a common measurement of the level of harmonic distortion present in power systems. THD is defined as the ratio of total harmonics to the value at fundamental frequency. THD_I and THD_V values can be calculated with using following equations, [20] .

Total current harmonic distortion can be expressed as,

$$THD_I = \frac{\sqrt{\sum_{n=2}^{\infty} I_n^2}}{I_1} \quad (1)$$

Total voltage harmonic distortion can be expressed as,

$$THD_V = \frac{\sqrt{\sum_{n=2}^{\infty} V_n^2}}{V_1} \quad (2)$$

Harmonic currents causes the voltage drop across the impedance of the electrical system. This voltage drops are added fundemantal voltage and this causes voltage distortion or an increase in effective voltage. One of the major effects of power system harmonics is to increase the current in the system. Non sinusoidal power system currents (harmonic currents) causes a distortion of voltage waveform.. Therefore, power quality distortion and the resonance risk in the system is increasing. Therefore, this study focused on the current harmonics. In the next section THD_I measurements locations and the THD_I measurement values are given.

2.2 Power systems harmonics measurement

Manisa's energy distribution system is supplied from different transformer centers. The input voltage of distribution system is 154 kV. 154 / 34.5 kV transformers are used to turn voltage value to 34.5 kV in secondary distribution bus. THD_I values was measured at the 34,5 kV side of the transformer. THD_I values taken from residential areas and industrial zones are used in the ANN model to harmonic estimation. Single-line diagram of the distribution system is given in Figure 1. Residential areas THD_I measurements are given in first section, industrial zones THD_I measurements are given in second section .

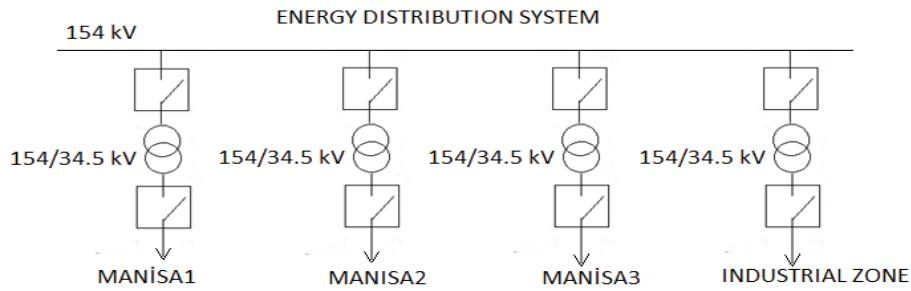


Figure 1. Single-line diagram of the energy distribution system

2.2.1 *Harmonic Measurements in Residential Areas.* Single-line diagram of the Residential Areas distribution system is given in Figure 2. Six different power line are connected to Manisa1 distribution line. Main transformer transforms the voltage from 154 kV to 34,5 kV, secondary distribution transformer transforms the voltage from 34,5 kV to 400 V. THD_I values was measured at the 34.5 kV side of the transformer. These measured THD_I values have been used by ANN for training and testing purposes. The load current and THD_I measurements were recorded 24 hour period.

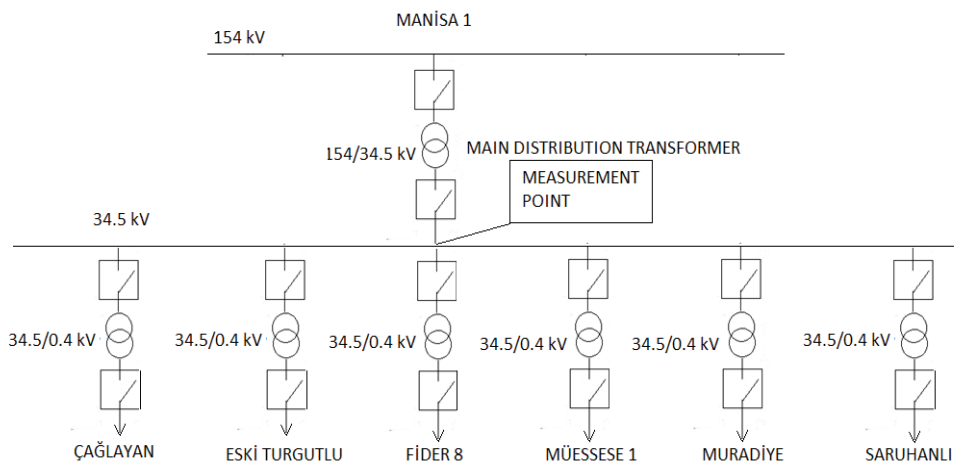


Figure 2. Single-line diagram of the Residential Areas energy distribution system

2.2.2 *Harmonic Measurements in Industrial Zone.* Single-line diagram of the Industrial zone distribution system is given in Figure 3. Four different power line are connected to Industrial zone distribution line. Main transformer transforms the voltage from 154 kV to 34,5 kV, secondary distribution transformer transforms the voltage from 34,5 kV to 400 V. THD_I values was measured at the 34,5 kV side of the transformer. These measured THD_I values have been used by ANN for training and testing purposes. The load current and THD_I measurements were recorded 24 hour period.

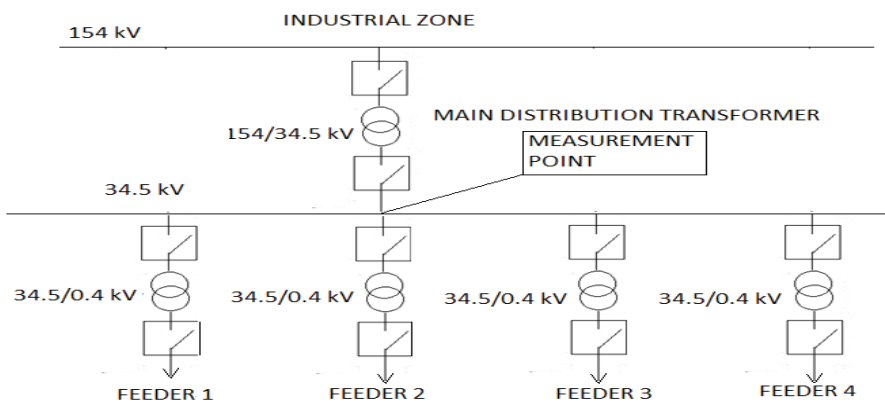


Figure 3. Single-line diagram of the Industrial zone energy distribution system

3. SIMULATION

The backpropagation algorithm is used in layered feed-forward ANNs. This means that the artificial neurons are organized in layers, and send their signals “forward”, and then the errors are propagated backwards. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. There may be one or more intermediate hidden layers. Each layer of processing elements are connected to the lower layer all operating elements. Also, the same layer there is no connection between the processing elements. Layers except for the input layer is responsible for processing information. [21]. Artificial neural networks are the method that commonly used in harmonic analysis. Modeling of non-linear loads, learning and the ability to generalize and the adaptability for different problems are ANN s important features.

The most important features in the ANN’s that determine the behavior of neurons is the neuron activation function Artificial neural networks are divided into three parts.

1. Input Layer
2. Hidden Layer
3. Output Layer

The number of input neurons, hidden layer neurons and the output layer neurons are determined according to the the parameters to be estimated.

Input layer neurons receive the input data and send this data to the hidden layer neurons. ANN used for estimation purposes, the most important performance is the accuracy of estimation. According to Makridakis and others commonly used error values were determined by the following formulas. Y to show the value of the real observation, F to show the value generated by the model [22], [15],

Estimation Error

$$e=Y-F \tag{3}$$

the average error

$$\frac{1}{n} \sum_{i=1}^n e_i \tag{4}$$

can be written with the above formulas. The average absolute error

$$\frac{1}{n} \sum_{i=1}^n |e_i| \tag{5}$$

$$\text{The percentage error } \frac{e_t}{Y_t} * 100 \tag{6}$$

can be written with the above formulas.

ANN gives better results than linear model in estimations (Gonzales, 2000) [23]. Different programs are used for the ANN model. In this study, the MATLAB program was used for the ANN model. Tansig and purelin have been used as an element of neuronal activation in the program. The measurements from residential areas and industrial zones and THD_I estimation results are shown in table form in the next sub-section.

3.1 Residential areas

The residential areas load current measurements and THD_I values are shown table 1. Measured data between 1 and 14 were used to train the ANN, and the data between 15 and 25 were used to test the ANN. Measured Load Current (I_L) values were used as the input, the THD_I values were used as the output in the ANN model. THD_I values were estimated with this ANN model.

Table 1. Measured Load Current (I_L) and THD_I values

| MEASUREMENT NUMBER | LOAD CURRENT (IL) | MEASURED THDI | MEASUREMENT NUMBER | LOAD CURRENT (IL) | MEASURED THDI | MEASUREMENT NUMBER | LOAD CURRENT (IL) | MEASURED THDI |
|--------------------|-------------------|---------------|--------------------|-------------------|---------------|--------------------|-------------------|---------------|
| 1 | 370 | 8,4 | 10 | 465 | 6,8 | 19 | 477 | 5,8 |
| 2 | 330 | 8,7 | 11 | 470 | 6,6 | 20 | 448 | 6,8 |
| 3 | 270 | 11,2 | 12 | 480 | 6,7 | 21 | 410 | 7,8 |
| 4 | 250 | 12 | 13 | 490 | 6 | 22 | 360 | 9 |
| 5 | 240 | 13 | 14 | 475 | 5,8 | 23 | 260 | 12,2 |
| 9 | 447 | 7 | 18 | 440 | 7,1 | 24 | 300 | 11,8 |
| 6 | 220 | 18 | 15 | 238 | 13,8 | 25 | 380 | 8,4 |
| 7 | 290 | 11,8 | 16 | 465 | 6,2 | | | |
| 8 | 365 | 8,8 | 17 | 425 | 7 | | | |

As a result of experiment ANN with two hidden layers provide the appropriate values for THD_I estimation in Residential areas. Therefore, for the Residential areas ANN model with one input, two hidden layer, one output was used. 5 neurons were used in the hidden layer. ANN model reaches required error rate value in 227 iterations and gives the estimated value of THD_I . ANN model parameter values are shown in table 2.

Table 2. ANN model parameter values

| Parameter | Value |
|----------------------------------|----------|
| Number of Input layer neuron | 1 |
| Number of Input layer neuron | 1 |
| The number of Hidden Layers | 2 |
| The number of neurons in Layer 1 | 5 |
| The number of neurons in Layer 2 | 5 |
| Layer 1 activation function | tansig |
| Layer2 activation function | tansig |
| Maximum iteration number | 3500 |
| Error limit | 1.00e-15 |

ANN model has been shown in Figure 4. As seen from Figure 8, ANN model has 1-input, 2 hidden layers and one output layer . The neurons in the hidden layer of the ANN model has no connection between themselves and establish connections with the other layer neurons. THD_I estimation has been done with the proposed ANN model.

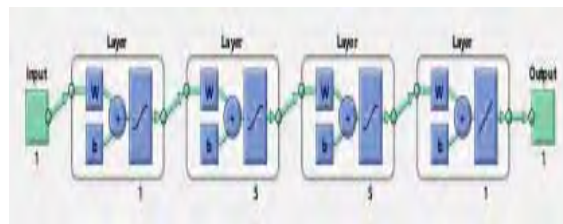


Figure 4. ANN model with One input layer, two hidden layer, one output layer.

Table 3. % error, measured THD_I , estimated THD_I values.

| MEASUREMENT NUMBER | MEASURED THD_I | ESTIMATED THD_I | % ERROR | MEASUREMENT NUMBER | MEASURED THD_I | ESTIMATED THD_I | % ERROR |
|--------------------|------------------|-------------------|---------|--------------------|------------------|-------------------|---------|
| 15 | 13,8 | 13,2 | 4,54 | 21 | 7,8 | 8,08 | 3,46 |
| 16 | 6,2 | 6,8 | 8,82 | 22 | 9 | 8,2 | 9,75 |
| 17 | 7 | 7,01 | 0,14 | 23 | 12,2 | 11,44 | 6,64 |
| 18 | 7,1 | 6,99 | 1,57 | 24 | 11,8 | 11,42 | 3,32 |
| 19 | 5,8 | 5,53 | 4,88 | 25 | 8,4 | 7,7 | 9,09 |
| 20 | 6,8 | 6,99 | 2,71 | | | | |

Table 3 gives the information about % error, the measured THD_I values, THD_I values estimated by ANN. The THD values in this table are used for testing the ANN. % Error value is expressed as percentage of the difference between the measured THD_I and estimated THD_I .

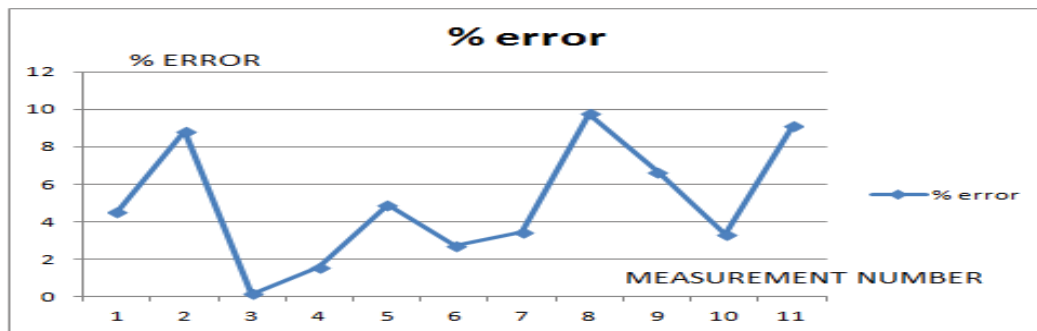


Figure 5. The % error between measured THD_I and ANN THD_I values.

Figure 5 shows the % error between THD_I and ANN THD_I values. First estimation error value is 4,54 % . Error value is then increased and reached to 8.4%. In third estimation error value has dropped to 0%. Estimated error value increased again in the 4th and 5th estimations, and reached to the value of 1.5% and 4.8% respectively. 6th estimate error value has decreased to 2.7%. In 7th and 8th estimates, error values have taken 3.4% and the highest error value 9.75% respectively. Then the error has taken the value of 6.6% in 9th estimation, 3.3% in 10th estimation, and 9% in 11th estimation. The highest estimation error in test data is 9.75%. in 9 numbered estimation. The lowest estimation error in test data is 0.14% in 3 numbered estimation. The average error for Residential areas is 4,99%. This indicates estimation has been done with the 95.01% accuracy.

3.2 Industrial Zone

Measured three-phase current (I_L) and THD_I values from Industrial Zone are shown in table 3. These values were used in order to train ANN. The I_L values were used as the input, the THD_I values were used as the output in the ANN model. THD_I values were estimated with this ANN model.

Table 4. Measured three-phase current (I_L) and THD_I values

| MEASUREMENT NUMBER | THE LOAD CURRENT PHASE 1 I_LN 1 | THE LOAD CURRENT PHASE 2 I_LN 2 | THE LOAD CURRENT PHASE 3 I_LN 3 | MEASURED THD_I 1 | MEASURED THD_I 2 | MEASURED THD_I 3 | MEASUREMENT NUMBER | THE LOAD CURRENT PHASE 1 I_LN 1 | THE LOAD CURRENT PHASE 2 I_LN 2 | THE LOAD CURRENT PHASE 3 I_LN 3 | MEASURED THD_I 1 | MEASURED THD_I 2 | MEASURED THD_I 3 |
|--------------------|---------------------------------|---------------------------------|---------------------------------|------------------|------------------|------------------|--------------------|---------------------------------|---------------------------------|---------------------------------|------------------|------------------|------------------|
| 1 | 215,4 | 218,4 | 220,1 | 8,2 | 7,6 | 7,9 | 13 | 234,2 | 232,6 | 236,5 | 5,9 | 5,8 | 5,6 |
| 2 | 202,2 | 205,4 | 206,2 | 8,4 | 8,4 | 8,1 | 14 | 232,7 | 233,9 | 235,7 | 6,3 | 6,1 | 6,1 |
| 3 | 210,5 | 211,6 | 216,1 | 8,8 | 8,8 | 8,6 | 15 | 237,7 | 236,5 | 235,8 | 6,4 | 6,2 | 6,3 |
| 4 | 222,6 | 224,2 | 227 | 7,9 | 7,7 | 7,8 | 16 | 238,5 | 237,8 | 241,1 | 5,3 | 5,3 | 5,1 |
| 5 | 207,4 | 204,7 | 210,3 | 9,7 | 10 | 9,5 | 17 | 213,7 | 213,1 | 216,7 | 7,1 | 7 | 6,8 |
| 6 | 199 | 199,2 | 201 | 9 | 9 | 8,5 | 18 | 205,9 | 205 | 207,2 | 7,6 | 7,3 | 7,4 |
| 7 | 203,3 | 205,8 | 207,8 | 8,3 | 8,3 | 7,8 | 19 | 191,1 | 189,8 | 190,8 | 8,5 | 8,3 | 8,1 |
| 8 | 191,6 | 189,5 | 192,2 | 9,5 | 9,4 | 8,8 | 20 | 193,5 | 197,2 | 197,4 | 8,1 | 7,5 | 7,4 |
| 9 | 241,8 | 240,6 | 244,5 | 7,2 | 7,1 | 6,9 | 21 | 214,8 | 216,6 | 220,2 | 8,2 | 8,1 | 7,8 |
| 10 | 251,1 | 252,6 | 255 | 6,5 | 6,4 | 6,4 | 22 | 219,7 | 219 | 224,2 | 7,4 | 7,2 | 7,2 |
| 11 | 249,8 | 249,9 | 253,4 | 5,9 | 5,9 | 5,7 | 23 | 215,1 | 215,4 | 220,6 | 8,2 | 8 | 7,9 |
| 12 | 252,6 | 252,6 | 257,3 | 7,3 | 7,3 | 7 | | | | | | | |

Table 4 indicates three phase load current and THD_I measurements. Measured data between 1 and 15 were used to train the ANN, and the measured data between 16 and 23 were used to test the ANN. As a result of experiment ANN with two hidden layers provide the appropriate values for THD_I estimation in Industrial zone. Therefore, for the Industrial zone ANN model with 3 input, 2 hidden layer, 3 output was used. 5 neurons were used in the hidden layer. ANN model, reaches required error rate value in 730 iterations and gives the estimated value of THD_I . ANN model parameter values are shown in table 5.

Table 5. ANN model parameter values

| Parameter | Value |
|----------------------------------|----------|
| Number of Input layer neuron | 3 |
| Number of Input layer neuron | 3 |
| The number of Hidden Layers | 2 |
| The number of neurons in Layer 1 | 5 |
| The number of neurons in Layer 2 | 5 |
| Layer 1 activation function | tansig |
| Layer2 activation function | tansig |
| Maximum iteration number | 1000 |
| Error limit | 1.00e-15 |

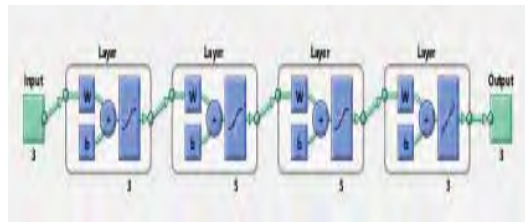


Figure 6. ANN model with 3 input layer, 2 hidden layer, 3 output layer.

ANN model has been shown in Figure 6 As seen from Figure 11, ANN model has 3 input, 2 hidden layers and 3 output layer . The neurons in the hidden layer of the ANN model has no connection between themselves and establish connections with the other layer neurons. THD_I estimation has been done with the proposed ANN model.

Table 6. % error, measured THD_I , estimated THD_I values.

| MEASUREMENT NUMBER | MEASURED THD_I1 | MEASURED THD_I2 | MEASURED THD_I3 | ESTIMATED THD_I1 | ESTIMATED THD_I2 | ESTIMATED THD_I3 | % ERROR THD_I1 | % ERROR THD_I2 | % ERROR THD_I3 |
|--------------------|-----------------|-----------------|-----------------|------------------|------------------|------------------|----------------|----------------|----------------|
| 16 | 5,3 | 5,3 | 5,1 | 5,3 | 5,3 | 5,3 | 0 | 0 | 3,7 |
| 17 | 7,1 | 7 | 6,8 | 7,4 | 7,3 | 7,1 | 4,1 | 4,1 | 4,2 |
| 18 | 7,6 | 7,3 | 7,4 | 8,3 | 8,2 | 7,7 | 8,4 | 11 | 3,8 |
| 19 | 8,5 | 8,3 | 8,1 | 8,4 | 8,2 | 7,6 | 1,2 | 1,2 | 6,5 |
| 20 | 8,1 | 7,5 | 7,4 | 7,8 | 8,7 | 7,7 | 3,8 | 14 | 3,9 |
| 21 | 8,2 | 8,1 | 7,8 | 8,1 | 7,9 | 7,88 | 1,8 | 2,8 | 1 |
| 22 | 7,4 | 7,2 | 7,2 | 7,8 | 7,9 | 7,5 | 5,1 | 8,8 | 4 |
| 23 | 8,2 | 8 | 7,9 | 8,2 | 7,5 | 8,3 | 0 | 6,6 | 4,8 |

Table 6 gives the information about % error, the measured THD_I values and THD_I values estimated by ANN. The THD_{I1} , THD_{I2} , THD_{I3} values in this table are used for testing the ANN. % Error value is expressed as percentage of the difference between the measured THD_I and estimated THD_I .

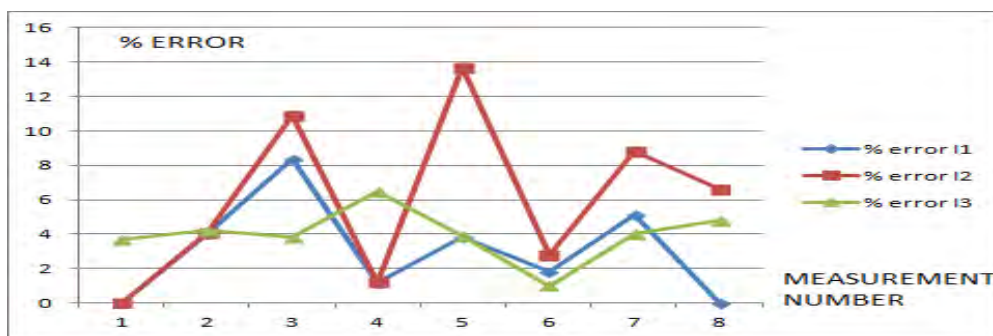


Figure 7. The % error between measured THD_I and ANN THD_I values for 3-phase system.

Figure 7 shows the % error between THD_I and ANN THD_I values for 3 phase system. First estimation error value for THD_{I1} is 0 % . Error values is then increased and reached to 4% and 8% . In 4th estimation error value has dropped to 1%. Then the error value has taken the values of 3.8%, 1.8% and 5.1% respectively. 8th estimation error value has taken a value of 0%.

First estimation error value for THD_{I2} is 0 % . Error values is then increased and reached to 4% and 11% . In 4th estimation error value has dropped to 1%. Error value, reached maximum value in the

fifth estimation and it has taken a value of 14 %. Then the error value has taken the values of 2.8%, 8.8% and 6.6% respectively.

First estimation error value for THD_1 is 3.7 % . Then the error value decreases and increases in very small range. Error value, reached maximum value in the 4th estimation and it has taken a value of 6.5 %. Then, the error has declined to 3.9 and 1 values respectively. Error values is then increased and reached to 4% and 4.8% values in 7th and 8th estimations.

For THD_1 the highest estimation error in test data is in the 3 numbered estimation with the value of 8.4% . The lowest estimation error in test data is 0% in 1 and 8 numbered estimations. The average error for THD_1 is 3,04. This indicates that the estimation has been done with the 97 % accuracy.

For THD_2 the highest estimation error in test data is in the 5 numbered estimation with the value of 8.7% . The lowest estimation error in test data is 0% in 1 numbered estimation. The average error for THD_2 is 6,01%. This indicates that the estimation has been done with the 94 % accuracy.

For THD_3 the highest estimation error in test data is in the 4 numbered estimation with the value of 6.5% . The lowest estimation error in test data is 1.01% in 6 numbered estimation. The average error for THD_3 is 3,98%. This indicates that the estimation has been done with the 96 % accuracy.

4. CONCLUSION

In this study, values may occur in the power distribution system were estimated by ANN model. Analysis and research by using ANN have indicated that ANN accurately estimates THD_1 values. With this proposed method, harmonics which is one of power quality problem can be estimated previously to improve power quality. Especially harmonic analysis and the estimation is very important in distribution system. Because of the harmonic distortion the losses and the various defects occur. Harmonics is adversely affects power quality because of this type of effect. THD_1 values are estimated with a high accuracy by ANN algorithm. The values measured from the energy distribution system have been used by ANN for training and testing purposes. THD_1 estimation by ANN model was determined close to real THD_1 value. This indicates that proposed method can be used to estimate THD_1 values. According to this estimated THD_1 value, the required filter parameters are determined and applied to a power distribution system. In this way, by reducing the harmonic values, a considerable increase in power quality can be provided. With the proposed ANN THD_1 estimation model, power quality equipment can be preselected while the energy distribution systems planning. ANN THD_1 estimation model, also be used in other systems with the purpose of THD_1 estimation.

5. REFERENCES

- [1] V.E. Wagner et al. , "Effects of harmonics on equipment," IEEE Trans. Power Del., vol. 8, no. 2, pp. 672-680, 1993.
- [2] R. Yacamini, "Power system harmonics: Part 3. Problems caused by distorted supplies," Power Engineering Journal, vol. 9, no. 5, pp. 233-238, 1995.
- [3] J. Arrillaga and N. R. Watson, Power System Harmonics.: John Wiley & Sons, 2003.
- [4] IEEE Guide for Harmonic Control and Reactive Compensation of Static Power Converters, ANSIIEEE Std. 519-1981, 1981.
- [5] IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems, IEEE Std. 519-1992, 1992.
- [6] Limits for Harmonic Current Emissions (Equipment Input Current Less than i6 A Per Phase), IEC Std. 61000-3-2, 2005.
- [7] Limitation of Emission of Harmonic Currents in Low-Voltage Power Supply Systems for Equipment with Rated Current Greater than i6 A, IEC Std. 61000-3-4, 1998.
- [8] U.Arumugam ,N.M. Nor,M.F. Abdullah "A Brief on Advances of Harmonic State Estimation Techniques in Power Systems." International Journal of information and Electronics Engineering Vol.1,No.3,November 2011
- [9] Sachin K. Jain and S. N. Singh, "Harmonics Estimation in Emerging Power System: Key Issues and Challenges," Electr. Power Syst. Res., vol.81, pp. 1754-1766,2011.
- [10] HuaOuyang,Jialin Wang " Power System Transient Signal Analysis based on Prony Algorithm and neural network," IEEE PES ISGT ASIA 2012 1569605083.

- [11] Wang Li,Wang Xiao-hua ,'' Application Of Neural Network In Power System Inter_Harmonic Dedection '' 978-0-7695-4360-4/10 2010 IEEE
- [12]Ibrahim El-Amin, Ihab Arafah '' Artificial Neural Network for Power System _Harmonic Estimation'' 0-7803-5105-3/98 1998 IEEE
- [13]Sachin K. Jain ,D.Saxena,S.N. Singh ''Adaptive Wavelet Neural Network Based Harmonic Estimation Single –Phase Systems'' 978-1-4673-6008-1 2012 IEEE
- [14] Yalçınöz T. , Erdem S. , Eminoğlu U. , "Yapay sinir ağları ile Niğde Bölgesinin elektrik yük tahmini"
- [15] Chen Ying,Lin Qingsheng "New Research on Harmonic Dedection Based Neural for Power System'' 978-0-7695-3589-4/09 2009 IEEE
- [16] H. C. Lin, "Intelligent neural network-based Dynamic power system harmonic detection," 0-7803-8610-8/04 2004 IEEE
- [17] Liu Qian-jin,Qin Si-shi, "A DFP- Neural Network Algorithm for Analysis of Power System Harmonics '' 978-1-4244-4813-5/10 2010 IEEE
- [18]Mario Oleskovicz,Marcelo A.A. Lima,Etienne Biassoto,Denis V. Coury '' Estimation of Harmonic Currents Injected by Nonlinear Loads for a Distorted Power Supply Scenario Using Artificial Neural Networks'' 978-1-1673-1943-0/12 2012 IEEE
- [19] Hsiung Cheng Lin, "Intelligent neural network-based fast power system harmonic detection," IEEE Trans. ind. Electron., vol. 54, no. 1, pp. 43-52,2007.
- [20]KOCATEPE C.'Elektrik tesislerinde harmonikler'', Birsen, 2.1-13,2003
- [21] Hamzaçebi C.,''Yapay Sinir Ağları'', Ekin, 4-104, 2011
- [22] Makridakis S., Wheelwright,S.C.,Mc Gee,V., ''Forecasting:Methods and applications'' John Wiley & Sons, NY,3-476 1983.
- [23] Gonzales S.'''Neural Netorks for Forecasting:A Complemantry Approach to Linear Regression Methods ,Working Paper 2000-07 ,Canada,26-33 (2000)

Comparison with different models of bending stress analysis of the cantilever beams under different profile section, materials and loads

Jabbar GATTMAH¹, Murat Tolga OZKAN², Ihsan TOKTAS³, Eylul DEMIR⁴

^{1,3,4}Engineering and Natural Science Faculty, Mechanical Engineering Department, University of Yıldırım Beyazıt, Ankara, Turkey

²Technology Faculty, Industrial Design Engineering Department, University of Gazi, Ankara, Turkey

¹*msc_jgj_katma7@yahoo.com*, ²*tozkan@gazi.edu.tr*, ³*itoktas@ybu.edu.tr*, ⁴*edemir@ybu.edu.tr*

ABSTRACT

In this study, the constraints and requirements for the analysis are determined during the cantilever beam model by using 7 different lengths and 4 different profiles with 7 different loads. The materials used in this study are 61 types. The bending stresses on a cantilever beam model are calculated. The results of empirical model are compared with the results of Regression Analysis (REGA), Finite Element Analysis (FEA), Artificial Neural Network (ANN) Model. In this work, four different methods are used to show maximum deflection at the bending stresses and tested the accuracy of empirical model. The ANN model gives the best results. The Statistical error analysing methods are used to compare the empirical results and ANN predictions. The Levenberg-Marquardt (LM) learning algorithm is used the ANN with 9+11+15 three hidden layers to produce Absolute Fraction of variance R^2 values approximately to 1 and the mean % errors and Root Mean Square Error (RMSE) values are found to be very low. It can be used ANN model instead of empirical model. The ANN with an acceptable accuracy is used to determine maximum deflection at the bending stresses. This study shows that ANNs can be used as an alternative method to determine maximum deflection of bending stresses on beams.

Keywords – *beam, bending stress, regression analysis, finite element analysis, artificial neural network*

1. INTRODUCTION

Large development that happens in computer technology leads to the use modelling techniques in scientific research extensively in present. Computational based approaches are use during Artificial neural networks [1]. The name “artificial neural networks” describes the hardware or software simulators, which are realizing semiparallel data processing. They are built from many mutually joined neurons and they imitate the work of biological brain structures [2]. To achieve good performance, they employ a massive interconnection of simple computing cells referred to as ‘Neurons’ or ‘processing units’. Hence a neural network viewed as an adaptive machine can be defined as A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making it available for use [3,4].

The employment of mechanical models in statistics has a long history [5]. Mechanical contrivances of all sorts have been brought to bear on a plethora of statistical concepts [6]. In particular, mechanical models for linear regression have been proposed [7]. Although of primary utility as conceptual devices [8], machines have actually been built for mechanically constructing Regression lines.

Numerical methods provide a general tool to analyze arbitrary geometries and loading condition. Among the numerical methods, Finite Element Analysis (FEA) has been extensively used with success; however, this kind of analysis requires the generation of a large set of data in

order to obtain reasonably accurate results and consumes large investment in engineering time and computer resources [9]. FEA gives accurate solutions through the using of engineering analysis and depends on FEA in the applications of mechanical engineering by using developed computer program [10], because direct connection with computers through mathematic equations and steps of running. To built cantilever beam model there are many software programs (Mark, Msc, ANSYS, Abaqus) are used and stored in the computers, in this study ANSYS program is used, and analyze this program which does not take much time for analyzing accomplishment [11].

This study has a new perspective to determination maximum deflection of a cantilever beam, and contains four different methods are emprical results, Regression Analysis (REGA), FEA and Artificial Neural Network (ANN) model. This study has put in to reveal the deviation of emprical model, FEA, ANN model and REGA. All models have been compared with each other.

2. MATERIAL AND METHOD

There are 61 types of materials are used to built cantilever beam model with use four methods are Emprical, REGA, FEA and ANN Model.

2.1 Determination of material properties

Axial force is applied to a cantilever beam, as shown in Figure 1. The Lenght of the cantilever beam is L and the applied axial force is F .

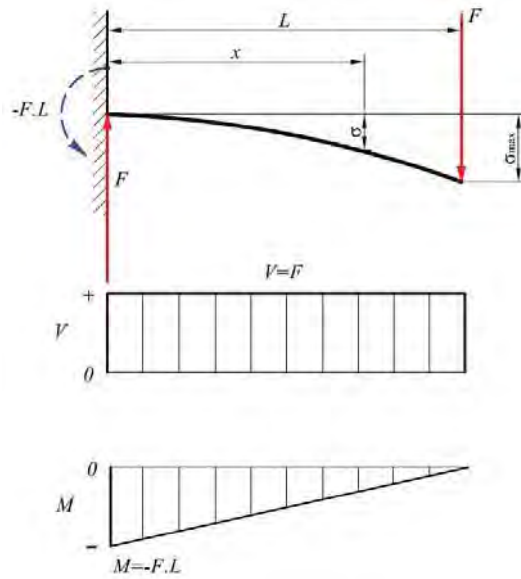


Figure 1. Model of cantilever beam and direction of forces

Eqs (1-4) are used for derivation of deflection (δ). Eq (5) is used to calculate maximum deflection δ_{max} where M is bending moment, E is modulus of elasticity and I is moment of inertia:

$$M_{AB} = (F.X) - M \quad (1)$$

$$y'' = \delta = \frac{1}{EI} \int [(F.X) - M] dx \quad (2)$$

$$\delta = \frac{1}{EI} \left[\frac{FX^2}{2} - M.X + C_1 \right] \quad (3)$$

$$y' = \delta_{max} = \frac{1}{EI} \left[\frac{FX^3}{6} - \frac{MX^2}{2} + C_1X + C_2 \right] \quad (4)$$

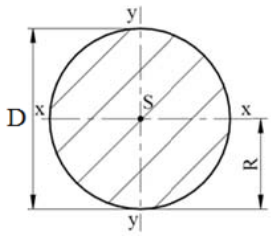
Where $C_1, C_2 = 0$ and $X = L$ we obtain Eq (5)

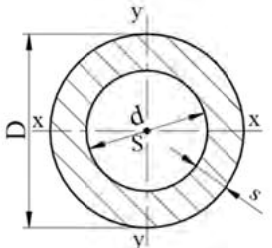
$$\delta_{max} = \frac{1}{EI} \left[\frac{FL^3}{6} - \frac{ML^2}{2} \right] \quad (5)$$

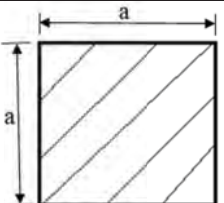
The four different profiles of cantilever beam which are used in this study depend on cross section are of cantilever beam: first profile is circular and depend on diameter of cantilever beam, the second profile is cylindrical depend on outer diameter and inner diameter of cantilever beam, square profile is the third type and depend on equal width of cantilever beam and the fourth profile is rectangular depend on height and width of cantilever beam. These profiles are built with different dimensions Table 1 and after that the axial forces are applied with different values to obtain the results.

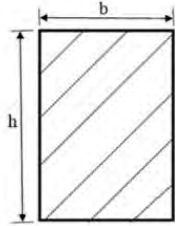
All of the profiles for cantilever beam can be seen at Table 1:

Table 1. Profiles for cantilever beam [12]

| Profile Types (Circular profile) | Length (L-mm) | Diameter (D-mm) | Force (F-N) |
|---|------------------|--------------------|----------------|
|  | 50 | 30 | 1000 |
| | 100 | 50 | 1500 |
| | 200 | 100 | 2000 |
| | 300 | 150 | 5000 |
| | 500 | 200 | 10000 |
| | 750 | 240 | 15000 |
| | 1000 | 280 | 20000 |
| | 1000 | 300 | 20000 |

| Profile Types (Cylindrical profile) | Length (L-mm) | Diameter (D-mm) | Thickness (S-mm) | Force (F-N) |
|---|------------------|--------------------|---------------------|----------------|
|  | 50 | 80 | 10 | 1000 |
| | 100 | 110 | 18 | 1500 |
| | 200 | 160 | 24 | 2000 |
| | 300 | 190 | 18 | 5000 |
| | 500 | 200 | 30 | 10000 |
| | 750 | 240 | 30 | 15000 |
| | 1000 | 320 | 40 | 20000 |
| | 1000 | 400 | 40 | 20000 |

| Profile Types (Square profile) | Length (L-mm) | Width (a-mm) | Force (F-N) |
|---|------------------|-----------------|----------------|
|  | 50 | 2 | 1000 |
| | 100 | 4.5 | 1500 |
| | 200 | 8 | 2000 |
| | 300 | 14 | 5000 |
| | 500 | 16 | 10000 |
| | 750 | 25 | 15000 |
| | 1000 | 45 | 20000 |
| | 1000 | 65 | 20000 |

| Profile Types (Rectangular profile) | Length (<i>L</i> -mm) | Width (<i>b</i> -mm) | Height (<i>h</i> -mm) | Force (<i>F</i> -N) |
|---|---------------------------|--------------------------|---------------------------|-------------------------|
|  | 50 | 2 | 3 | 1000 |
| | 100 | 4.5 | 6.75 | 1500 |
| | 200 | 8 | 12 | 2000 |
| | 300 | 14 | 21 | 5000 |
| | 500 | 16 | 24 | 10000 |
| | 750 | 25 | 37.5 | 15000 |
| | 1000 | 45 | 67.5 | 20000 |
| | | | 65 | 97.5 |

2.2 Regression Analysis

Regression analysis is performed for the modeling and analysis of several variables where there is a relationship between a dependent variable and one or more independent variables. The correlations between the parameters applied force and dimensions of cantilever beam (length, outer diameter, inner diameter, width and height), the output of maximum deflections are obtained by multiple linear regression. Through the backward elimination process, the final quadratic models of the response equation in terms of actual deflection are presented as follows (Eq. 6):

$$Deflection_{REGA} = (1E - 08x^2) + (8E - 08x) + (6E - 06) \quad (6)$$

The coefficient of determination (R^2) defines the correlation between experimental and predicted results. The differences between the real responses which were measured after the experiments and the estimated responses that were calculated with the above equation is given. This REGA was presented [13]. This model has $R^2 = 0.986425125$.

ANOVA helps in formally testing the significance of all main factors by comparing the mean square against an estimate of the experimental errors at specific confidence levels [13]. The optimal combination of process parameters was predicted by both *S/N* and ANOVA analysis. The ANOVA results obtained for maximum deflections according to parameters at the confidence level of 95% was performed.

The ANOVA table is a widely used model which tests the significance and interactions of the factors and the suitability of the model [14]. The F-Test is calculated in order to show by what amount the test results are affected by the control factors. In the ANOVA table, when “Prob > F” is less than 0.05, the control factors and their interactions are significant [15].

2.3 Finite Element Analysis (FEA)

Cantilever beam with circular, cylindrical, square and rectangular profile are used 3D to built the models using Finite Element Analysis (FEA) method. Study of parameters are different for four profiles. The first step of the model determination of mesh optimization (Fig. 2). This optimization has continued until the FEA results. After the determination of mesh model, different axial forces applied the on varried cantilever beam that have different dimensions sistematically. Then the values of maximum deflections for Cantilever models are collected and classified according to model dimesions and applied force. Then anaysis of models for obtaining the results in Fig. 3 to compared with the empirical results

The elastic stress and strain fields of FEA are systematically investigated using 3D FEA method. It is found that the maximum stress and strain do always occur on the fixed part of cantilever beam. The maximum deflections of the cantilever beam model are different and depend on type of material, these deflections increase with increasing of length and decrease with increasing the cross section area. The contour of deflection is showed that minimum of deflection is being on fixid part and equal zero then it starts to increase gradually until reaches to maximum value at point of force *F*.

A parametric FEA model has been developed for determination of deflections in a circular cantilever beam. First step was mesh optimisation for workpiece. Lenght, diameter and applied force have been choosen as analysis parameters from Table 1. Different mesh types and number of elements have been checked to optimise mesh for the model. After that different bending forces have been applied the model. These model results have been compared with analytical solution, many analysis have been performed using FEA model using ANSYS software. Empirical results

have been collected and classified. FEA results have been compared with the others models. These models were empirical model, REGA, ANN model and the FEA model.

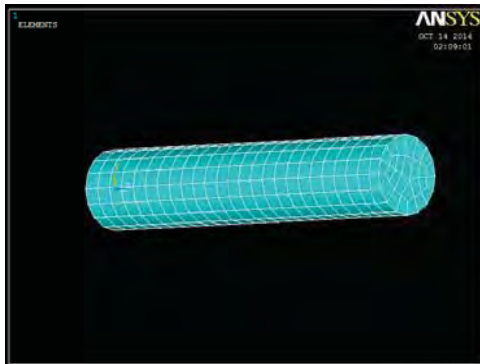


Figure 2. Mesh elements

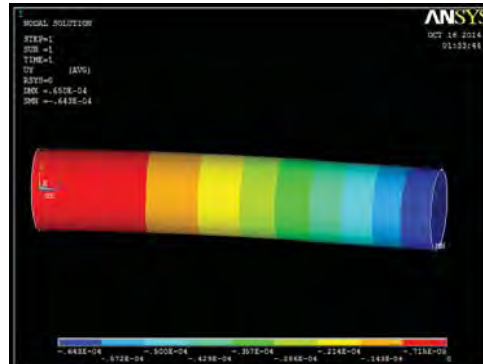


Figure 3. Deflections of beam

2.4 Artificial Neural Network (ANN) Model

A neuron is the basic element of neural networks and depending on its duties its shape and size may vary. Analyzing a neuron in terms of its activities is important, since understanding the way it works also helps us to construct the ANNs. An ANN may be seen as a black box which contains hierarchical sets of neurons (e.g., processing elements) producing outputs for certain inputs [16].

Each processing element consists of data collection, processing the data and sending the results to the relevant consequent element. The whole process may be viewed in terms of the inputs, weights, the summation function, and the activation function in Fig. 4 [17].

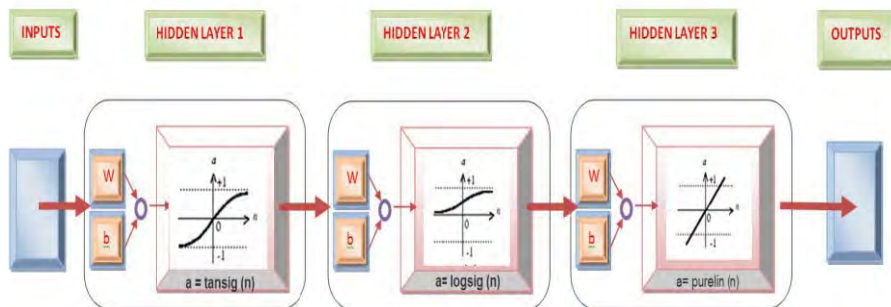


Figure 4. Basic artificial neural network model [16]

According to the figure 4, we have the following:

- I. The inputs, are the activity of collecting data from the relevant sources.
- II. The weights, control the effects of the inputs on the neuron. In other words, an ANN saves its information over its links and each link has a weight. These weights are constantly varied while trying to optimize the relation in between the inputs and outputs.
- III. Summation function, is to calculate of the net input readings from the processing elements.
- IV. Transfer (activation) function, determines the output of the neuron by accepting the net input provided by the summation function. There are several transfer functions f like summation function. Depending on the nature of the problem, the determination of transfer and summation function are made. A transfer function generally consists of algebraic equations of linear or nonlinear form [18]. The use of a nonlinear transfer function makes a network capable of storing nonlinear relationships between the input and the output. A commonly used function is sigmoid function because it is self-limiting and has a simple derivative. An advantage of this function is that the output cannot grow infinitely large or small [19].
- V. Outputs, accept the results of the transfer function and present them either to the relevant processing element or to the outside of the network.

The functioning of ANNs depends on their physical structure. An ANN may be regarded as a directed graph containing a summation function, a transfer function, its structure, and the learning rule used in it. The processing elements have links inbetween them forming a layer of networks. A neural network usually consists of an input layer, a number of hidden layers, and an output layer [17].

In this study, the training and test data have been prepared using empirical patterns. The data were obtained according to study parameters that has 83000 lines. Among them, 209 patterns each profile shape have been randomly selected and used as the test data. About 30% data were used as test data, 70 % data were used training data. Length, width, thickness, diameter and applied force were used as input-layer, LM algorithm and MLP (Multi Layer Perception) were used in the ANN model. Deflection was used as output-layer of the ANNs. In the ANN model, tansig, logsig and purelin transfer functions have been used and expressed as follows in Eqs (7-10):

$$NET_i = \sum w_{ij} \cdot x_j + w_{bi} \quad (7)$$

$$a = \text{tansig}(n) = \frac{2}{(1+e^{-2n})} - 1 \quad (8)$$

$$a = \text{logsig}(n) = \frac{1}{(1+e^{-n})} \quad (9)$$

$$a = \text{purelin}(n) \quad (10)$$

where NET is the weighted sum of the input. Input and output values (o) are normalized between 0 and 1. n : Number of processing elements in the previous layer.

Generally, there are three different learning strategies. Firstly, the trainer may tell the network what it should learn (supervised learning), secondly, the trainer may indicate whether or not the output is correct without telling what the network should learn (reinforcement learning), and finally, the network learns without any intervention of the trainer (unsupervised learning). The learning set consists of the inputs and the outputs used in training the network. The required outputs take place in this set in the case of supervised learning, while in other cases, they are not found in it [17]. In our case, we have used supervised learning approach.

Since the number of neurons found in the input and output layers are known, the best performance of the network with the number of hidden layers is determined using trial error method. Using limited number of neurons with limited number of hidden layers causes lesser learning, while increasing these numbers too much, decreases the speed of learning, and in some cases prevents the learning entirely. Usually, an algorithm is used for the learning process, this algorithm determines the weights. There are various learning methods using these strategies [17]. The back propagation learning algorithm has been used with Scaled conjugate gradient learning algorithm (SCG) and Levenberg marquardt learning algorithm (LM) versions at the training and testing stages of the networks [20]. The computer program has been developed under MATLAB [21]. In the first step of the training, a determination of the learning algorithms is made. The number of hidden layers and the number of neurons for each hidden layer are determined. Then, the number of iterations are entered by the user, and the training starts. The training continues either to the end of the iterations or reaching the target level of errors. Fig. 5 illustrates the ANN predictions against the empirical results.

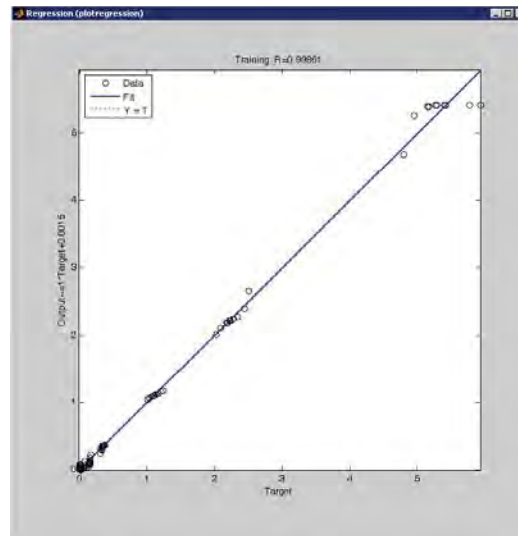


Figure 5. The ANN predictions against the empirical-based results.

3. TESTING THE ACCURACY OF REGRESSION ANALYSIS AND ANN-BASED

In order to understand whether a multiple regression analysis or an ANN is making good predictions, the test data which has never been presented to the network is used and the results are checked at this stage. The statistical methods of Root Mean Square Error (*RMSE*), Absolute fraction of variance (R^2), and Mean Error Percentage (*MEP*) values have been used for making comparisons [16,22,23]. The same data obtained from the regression analysis is used to determine the mentioned values.

These values are determined by the following Eqs (11-13):

$$RMSE = \left((1/p) \sum_j |t_j - o_j|^2 \right)^{1/2} \quad (11)$$

$$R^2 = 1 - \left(\frac{\sum_j (t_j - o_j)^2}{\sum_j (o_j)^2} \right) \quad (12)$$

$$MEP = \frac{\sum_j \left(\frac{t_j - o_j}{t_j} \times 100 \right)}{p} \quad (13)$$

where t is the target value, o the output, and p the number of samples.

Using the trial error method, the structure of the network (i.e. the number of neurons and hidden layers) is altered and the training operation is repeated. To be able to get accurate results we have used two hidden layers and for each hidden layer we have changed the number of neurons used at each hidden layer (e.g. from 5 to 150) to get the best network in terms of the statistical errors that it provides.

4. RESULTS AND DISCUSSION

In this study we have composed the empirical solution and network predicted deflection results for the cantilever beam statistical error analysing methods. As presented in Table 1 the statistical error levels for both training and testing data sets are evaluated. As the Table 2 illustrates the network with three hidden layers of 4+9+11+15+1 neurons at each layer has provided the best results (Fig. 6).

Following, the ANN model as illustrated in Fig.6 set up using 4 neurons for the input layer and with 9+11+15 processing elements at three hidden layers and finally 1 neuron are used at the output layer. The representation of knowledge is accomplished by the weights inbetween the layers.

In terms of the statistical error analysis methods, using Levenberg-Marquardt (LM) learning algorithm technique for maximum deflection, the mean error value for the training data set is % 0.199322746 while for the testing data it becomes % 0.091683771 (Table 2).

Analysis results have been collected and classified according to the 7 different lengths of the cantilever beam (L), four different profiles with 61 type of material and the applied 7 different axial forces (F). FEA results have been collected according to circular profile, 7 different lengths of the cantilever beam (L) with 11 modules of elastisty *and* applied 7 different axial forces (F), maximum deflection results have been compared with the others models, these models were Emprical model, Regression Analysis (REGA), Artificial Neural Network (ANN) model.

This optimization has continued until the FEA results and emprical solutions close to each other. Maximum deflections on the circular cantilever beam are collected and classified according to model dimensions and applied force. Then FEA have been performed and compared with the emprical results.

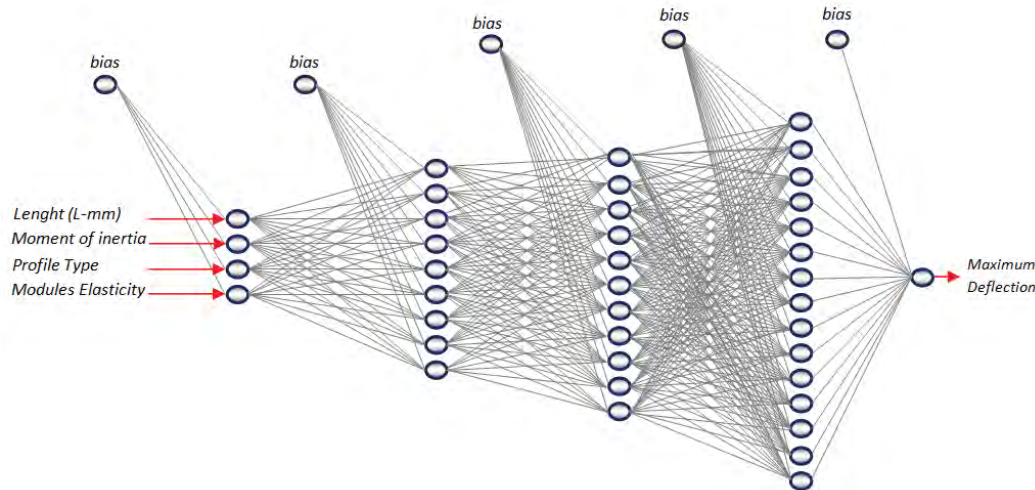


Figure 6. ANN architecture with 9+11+15 processing elements at three hidden layers.

Table 2. Comparison of ANN model, REGA model and FEA with emprical solution using statistical error analysing methods for the cantilever beam model

| | | Absolute Fraction of Variance (R^2) | Root Mean Square Error (RMSE) | Mean Error Percentage (MEP %) |
|---|------|---|-------------------------------|-------------------------------|
| Deflections_{ANN} | Test | 0.999958163 | $1.28318.10^{-13}$ | 0.091683771 |
| Deflections_{Regression} | - | 0.986425125 | $2.98194.10^{-8}$ | 0.942807101 |
| Deflections_{FEA} | - | 0.999894866 | $2.01916.10^{-8}$ | 0.124531710 |

Fig. 7 shows relation between deflections (δ) and Modulus of Elasticity (E) for circular profile type ($L=50\text{mm}$, $A=757\text{mm}^2$). This graphs emphasis to variation of deflection. When Modulus of Elasticity (E) increases, deflection decreases.

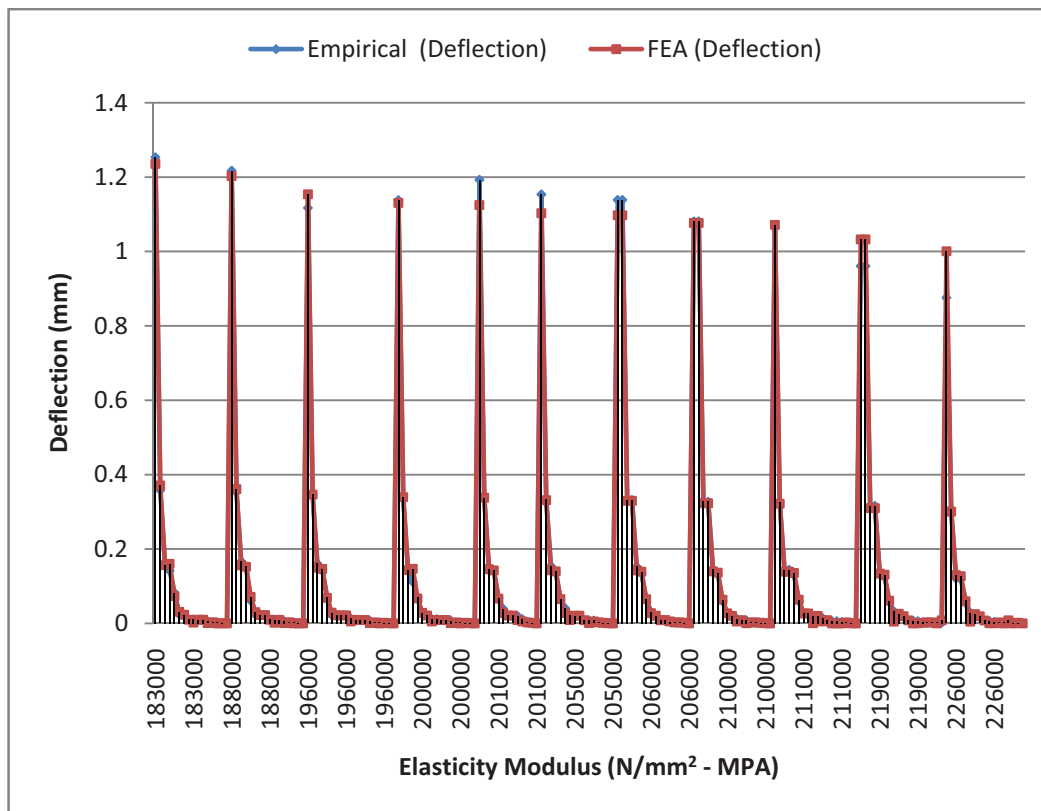


Figure 7. Dependence of Displacement on Modulus of Elasticity Diagram (circular profile type, $L=50\text{mm}$, $A=757\text{mm}^2$)

5. CONCLUSIONS

In this study, the application of regression analysis, ANN and FEA on the empirical deflection values were compared and discussed. This study revealed some inaccuracy in the empirical model and FEA. The developed models, which are limited by their boundary conditions, are compared in terms of prediction accuracy to deflection. For a long time, modeling techniques have been developed for the prediction of deflection. From the results of this study, conclusions are drawn. It is concluded that ANNs can be used as an alternative method for the determination of deflection for a cantilever beam.

6. REFERENCES

- [1] Sha W., and Edwards K. L., "The use of artificial neural networks in materials science based research," in *Materials & Design*, vol. 28, Issue 6, 2007, pp. 1747-1752.
- [2] Tadeusiewicz R., "Elementary introduction for neural networks techniques with sample application," in Academic Publishing House PLJ, Warsaw, 1998, in Polish.
- [3] Panchanand J. H. E., "Novel Artificial Neural Network Application for Prediction of Inverse Kinematics of Manipulator," Department of Mechanical Engineering, Institute of Technology, Master Thesis, 2009.
- [4] Shtub A., and Versano R., "Estimating the cost of steel pipe bending, a comparison between neural networks and regression analysis," in *International Journal of Production Economics*, vol. 62, Issue 3, 1999, pp. 201-207.
- [5] Farebrother R. W., "Fitting Linear Relationships, A History of the Calculus of Observation," Springer, New York, 1999, pp. 1750-1900.
- [6] Farebrother R. W., "Visualizing Statistical Models and Concepts", Marcel Dekker, New York, 2002.

- [7] Farebrother R. W., "Mechanical Representations of the L1 and L2 Estimation Problems, Statistical Data Analysis Based on the L1-norm and Related Methods", Y. Dodge (Editor), Elsevier North-Holland, New York, 1987, 455-464.
- [8] Boyd J. N. and Raychowdhury P. N., "A Statistical Result Derived from Mechanical Equilibrium," in Virginia Journal of Science, vol. 44 (4), 1993, pp. 341-344.
- [9] AL-Momani, E., and Rawabdeh, I., "An Application Of Finite Element Method And Design Of Experiments In The Optimization Of Sheet Metal Blanking Process," in Joudan Journal of Mechanical and Industrial Engineering, vol. 2, Number 1, 2008, pp. 53-63.
- [10] Hinton. E., and Owen, D. R. J, "Finite Element Programming", Academic Press, London, 1983.
- [11] Gattmah J, K. J., "Calculation of Relative Extrusion Pressure for Circular section by Local Coordinate system by Using Finite Element Method," in Diyala Journal of Engineering Sciences, vol. 3, no: 2, 2010, pp. 80-96.
- [12] Mukavemet ve Malzeme (2009) Citation Reference [Online]. Available: [http:// www.tasarimveimalat.com/mukavemet.pdf](http://www.tasarimveimalat.com/mukavemet.pdf)
- [13] Ross P. J., "Taguchi Technicues for quality engineering", Mc Graw Hill International Edition, Singapur 1996.
- [14] Bas D., and Boyaci I., "Modeling and optimisation. II. Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction," in J. Food Eng., vol. 78, 2007, pp. 846-854.
- [15] Ko-Ta C., "Modeling and analysis of the effects of machining parameters on the performance characteristics in the EDM process of Al₂O₃+TiC mixed ceramic," in The International Journal of Advanced Manufacturing Technology, vol. 37, 2008, pp. 523-533.
- [16] Toktaş İ., and Başak H., "Chain Gears Design Using Artificial Neural Networks," in Computer Applications in Engineering Education, vol. 20, issue 1, 2012, pp. 38-44.
- [17] Oztemel E., "Integrating expert systems and neural networks for intelligent on-line statistical process control," Ph.D. thesis, School of Electrical, Electronic and Systems Engineering, University of Wales, Cardiff, Wales, UK, December 1992.
- [18] Hollingworth N., and Hills D. A., "Theoretical efficiency of a cranked link chain drive," in Proc Inst Mech Eng Part C, 200, 1986, pp. 375-377.
- [19] Massie D. D., "Neural network fundamentals for scientists and engineers," in Proceedings of the International Congress on Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems and Processes (ECOS '01), Istanbul, Turkey, 2001, pp. 123-128.
- [20] Hagan M. T. and Demuth H. B., "Neural Network Design", PWS, Boston, Mass, USA, 1996, pp. 1-29.
- [21] MATLAB 6.5 (Release 13), "The language of technical computing," The MathWorks, Inc., Natick, MA, 2002.
- [22] Ozkan M.T., "Experimental and artificial neural network study of heat formation values of drilling&boring operations on Al 7075 T6 workpiece," in Indian Journal of Engineering & Materials Science, vol. 20 (4), 2013, pp. 259-268.
- [23] Ozkan M.T., Eldem C., and Sahin I., "Determination of Notch Factor for Shafts Under Torsional Stress by the Help of Artificial Neural Networks," in Materiali in tehnologije / Materials and Technology, MTAEC9, vol. 48 (1), 2014, pp. 81-90.

Comparison with different models of tensile and compressive stress analysis on a cantilever beam model

Eylul DEMİR¹, Ihsan TOKTAS², Murat Tolga OZKAN³, Jabbar GATTMAH⁴

^{1,2,4}Engineering and Natural Science Faculty, Mechanical Engineering Department, University of Yıldırım Beyazıt, Ankara, Turkey

³Technology Faculty, Industrial Design Engineering Department, University of Gazi, Ankara, Turkey

¹edemir@ybu.edu.tr, ²itoktas@ybu.edu.tr, ³tozkan@gazi.edu.tr, ⁴msc_jgj_katma7@yahoo.com

ABSTRACT

In this study, the constraints and requirements for the analysis of the cantilever beam model are determined. According to 7 different lengths, 6 different profiles, 7 different cross sectional areas, 6 different loads and 60 different materials, the tensile and compressive stresses on a Cantilever Beam Model are investigated. The empirical model is compared with the Regression Analysis (REGA), Finite Element Analysis (FEA), Artificial Neural Network (ANN) Model. The tensile and compressive stresses are modelled with 4 different methods and tested the accuracy of empirical model. The best results are obtained using ANN model. The empirical results and ANN predictions are compared by using statistical error analysing methods. The ANN with 9+9+11+17 processing elements (i, j) at four hidden layers using Levenberg-Marquardt (LM) model produce Absolute Fraction of Variance (R^2) values approximately to 1 and the mean % errors and (RMSE) values are found to be very low. ANN model can be used instead of empirical model. The tensile and compressive stresses are determined by the ANN with an acceptable accuracy. It is concluded that ANNs can be used as an alternative method to determination of stresses on beams.

Keywords – beam, tensile and compressive stress, regression analysis, finite element analysis, artificial neural network

1. INTRODUCTION

The computer modelling techniques are used in scientific research extensively in present. Artificial neural networks has been introduced in order that computational based approaches are involved [1]. Artificial neural networks (ANN) are proposed with reference to the principles of human brain and it can produce solutions for complicated systems thanks to its features such as learning and parallel operation in a short time [2]. ANN are used widely at the present time and have applications such as classification, estimation, controlling systems, optimization, decision making. In recent years, use of ANN for modelling and estimation has increased [3-6].

Back propagation neural networks applications are examined for analysing of these problems: (1) bending analysis of elastoplastic beams, (2) elastoplastic plane stress problem, (3) estimation of fundamental vibration periods of real buildings, (4) detection of damage in a steel beam, (5) identification of loads applied to an elastoplastic beam. Regularization about neural network and application to estimation of concrete fatigue durability is discussed [7].

The flexural vibration in a cantilever beam which has a transverse surface crack is examined. The modal frequency parameters are calculated using fracture mechanics. The frequency values obtained are used to train a neural network to determine the crack location and depth [8]. For a castellated steel beam tending to buckling, load carrying capacity is calculated by nonlinear finite element method and failure condition is examined. An empirical model is also proposed to estimate critical load. Other alternative methods are traditional back-propagation (BP) neural

network and adaptive neuro-fuzzy inference system (ANFIS) are used. All of these methods are compared and ANFIS and neural network give the best accuracy [9].

A efficiency technique for calculating stress intensity factors is provided. Cantilever beam with finite width strip with central crack and a pin loaded circular hole with radial cracks are used. Consequently, numerical results and finite element analysis results have close agreement [10]. The large displacement behaviour of tapered cantilever beams subjected to end forces is studied by the finite element method. A finite element formulation derived and used. The numerical results show that the finite element formulation give the accurate results. The effect of the material inhomogeneity, taper ratio, and taper type on the large displacement are also pointed out [11]. Functionally graded cantilever beam is subjected to bending, shear force and pressure, respectively. Plane stress case is assumed and airy stress function is used. Elastic solutions for the beam are obtained. Effect of nonhomogeneous materials with different modulus on the elastic field in a cantilever beam is highlighted [12]. It is also available research about a proposed standard set of problems to test finite element accuracy, and finally it can be seen that the tests are able to display most of the parameters which affect finite element accuracy [13].

The stress intensity factor is calculated with an analytical method for cracked steel I-beams under both bending moment and axial load. The fatigue and fracture behavior of the steel I-beam are examined. These analytical results and the other results in the literature are compared [14].

This study has a new perspective to determination of tensile and compressive stresses of a cantilever beam. These study contains four different methods different from past studies that are empirical results, Regression Analysis (REGA), Finite Element Analysis (FEA) and Artificial Neural Network (ANN) model. This study has put in to reveal the deviation of empirical model, FEA, ANN model and REGA to calculate tensile and compressive stresses. All models have been compared with each other.

2. MATERIAL AND METHOD

Using empirical results, REGA, FEA and ANN Model solutions, the tensile and compressive stresses are calculated and compared with each other.

2.1 Determination of material properties

Axial load is applied to a cantilever beam, as shown in Fig.1. The Length of the cantilever beam is L and the applied axial load is F .

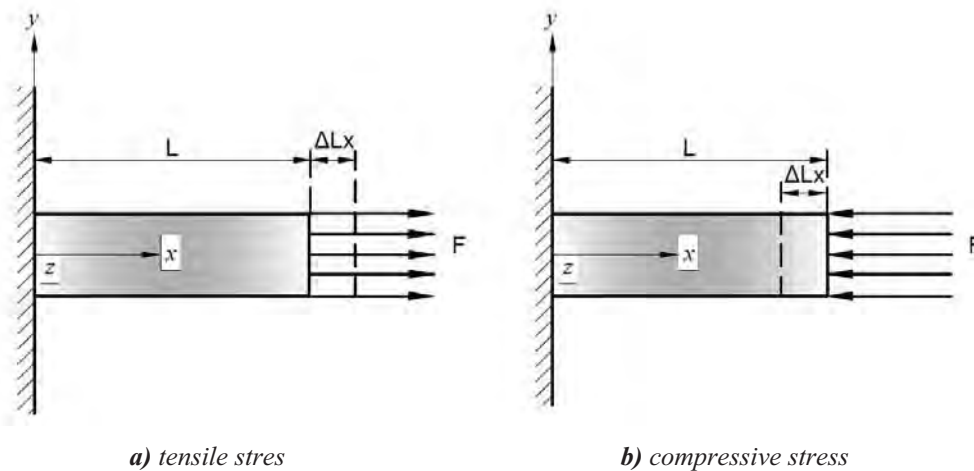


Figure 1. Model of cantilever beam and direction of loads

All of the profiles for cantilever beam can be seen at Table 1:

Table 1. Profiles for cantilever beam [15]

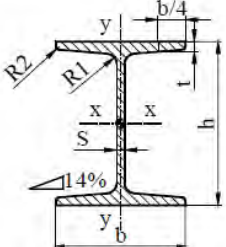
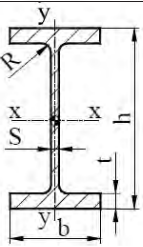
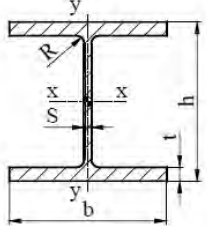
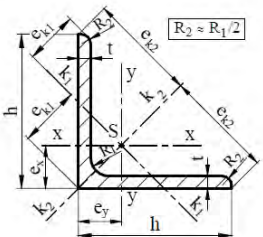
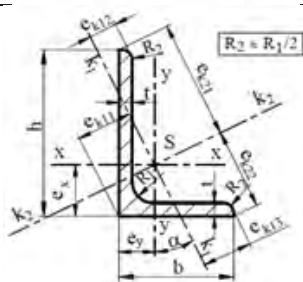
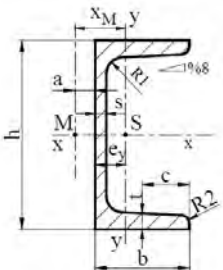
| Profile Types | Length (L-mm) | Area (A-mm ²) | Force (F-N) |
|---|---|--|---|
|  | 50 100 200 300 500 750 1000 | 757 1820 3340 5330 7770 10700 14700 | 1000 1500 2000 5000 10000 15000 20000 |
|  | 50 100 200 300 500 750 1000 | 764 1640 2850 4590 7270 11600 15600 | 1000 1500 2000 5000 10000 15000 20000 |
|  | 50 100 200 300 500 750 1000 | 2600 5430 9100 13100 17100 21800 27000 | 1000 1500 2000 5000 10000 15000 20000 |
|  | 50 100 200 300 500 750 1000 | 112 267 480 1227 2754 4606 7635 | 1000 1500 2000 5000 10000 15000 20000 |
|  | 50 100 200 300 500 750 1000 | 185 429 830 901 1150 2420 2920 | 1000 1500 2000 5000 10000 15000 20000 |
|  | 50 100 200 300 500 750 1000 | 221 492 1100 2400 4230 7480 9150 | 1000 1500 2000 5000 10000 15000 20000 |

Table 2 shows mechanical properties of materials used for cantilever beam:

Table 2. Mechanical properties of materials used for cantilever beam [15]

| DIN | W-Nr. | E, MPa | DIN | W-Nr. | E, MPa |
|---------------------|--------|--------|-----------------|--------|--------|
| NiCrTiAl F100 | 2.4952 | 183000 | 26NiCrMoV 14 | 1.6957 | 211000 |
| NiCo20Cr15 MoAlTi | 2.4634 | 188000 | 34CrMo4 | 1.722 | |
| X2Cr Ni 19 11 | 1.4306 | 196000 | 42CrMo4 | 1.7225 | |
| X6CrNi Ti 18 10 | 1.4541 | | 9SMnPb28 | 1.0718 | |
| X12Cr Ni 25 21 | 1.4845 | | GS-20Mn 5 N | 1.112 | |
| X6CrNi WNb 16 16 | 1.4945 | | 34Cr AlNi7 | 1.855 | |
| X12 Cr NiW Ti 16 13 | 1.4962 | | 30 CrMoV 9 | 1.7707 | |
| X8CrNiMoVNb 16 13 | 1.4988 | | 21CrMoV 5 7 | 1.7709 | |
| X2CrNiMo 17 13 2 | 1.4404 | | 23CrMo 5 | 1.7255 | |
| X6CrNiMoTi 17 12 2 | 1.4571 | | 13CrMo 4 4 | 1.7335 | |
| X5CrNi Ti 26 15 | 1.498 | | 15Mo 3 | 1.5415 | |
| NiCr13 Mo6Ti3 | 2.4662 | | 201000 | 22Mo 4 | |
| X12CrMoS 17 n | 1.4104 | 205000 | 23CrNi Mo 7 4 7 | 1.6749 | |
| X10CrAl 7 | 1.4713 | 206000 | 30Cr Mo Ni V 5 | 1.6946 | |
| X20CrNi 17 2 | 1.4057 | 210000 | 26Ni Cr MoV 11 | 1.6948 | |
| St50 | 1.005 | | 20CrMo NiV 4 7 | 1.6979 | |
| GS-C 25 | 1.0619 | | 21 Cr Mo V 5 11 | 1.807 | |
| ST33 | 1.0035 | | St35 BK | 1.0308 | |
| ST50-2 | 1.005 | | St35 GBK | 1.0308 | |
| St37-2 K | 1.0059 | | 10CrMo 9 10 | 1.738 | |
| St-37.0 N | 1.0244 | | St35.8 | 1.0305 | |
| St52-3 N | 1.057 | | Ck 75 H+A DIN | 1.1248 | |
| C 15 G | 1.0401 | | St 22 | 1.032 | |
| 15CrNi6 | 1.5919 | | G-X 22 Cr Mo V | 1.0619 | |
| C 35 | 1.0501 | 219000 | X20Cr13 | 1.4021 | |
| 20Mn 5 V | 1.1133 | | X20CrMoV 121 | 1.4922 | |
| Ck 35 | 1.1181 | | X12CrNi 21 21 | 1.4845 | |
| 30CrNiMo8 | 1.658 | | X22Cr MoV 121 | 1.4923 | |
| 34CrNiMo6 | 1.6582 | | X21 Cr MoV 12 | 1.4926 | |
| 28NiCrMoV 8 5 | 1.6932 | | NiCo 20 Co18 Ti | 2.4632 | |
| | | | | 226000 | |

In this study, tensile and compressive load effects on the profiles are investigated. Tensile and compressive loads were applied between 1000 N to 20000 N on the cantilever beam. Poisson ratio of steel beam is 0.3 and different profile types are selected. For different cross sectional areas (for example for I profile type; 757, 1820, 3340, 5330, 7770, 10700 and 14700 mm²) and different lengths (50, 100, 200, 300, 500, 750 and 1000 mm), it can be observed variation in stress with reference to results. The results are compared with Empirical Model, Finite Element Analysis (FEA), Regression Analysis (REGA) and Artificial Neural Network Model (ANN). Tensile stress can be calculated with Eq (1) and compressive stress can be obtained with Eq (2), strain and deformation are calculated with Eq (3-4) [15].

Tensile Stress:

$$\sigma_T = \frac{F}{A} \quad \frac{N}{mm^2} \quad (1)$$

A: cross-sectional area

F : load

Compressive Stress:

$$\sigma_c = \frac{-F}{A} = \frac{N}{mm^2} \quad (2)$$

$$\text{Normal Strain: } \varepsilon_x = \frac{\Delta L_x}{L} = \frac{\sigma_x}{E} = \frac{F}{E.A} \quad mm \quad (3)$$

E : Modulus of Elasticity or Modulus of Young,

Elongation $F > 0$ or

Contraction $F < 0$

Equating and solving for the deformation,

$$\Delta L_x = \frac{F.L}{E.A} \quad mm \quad (4)$$

Variation in displacement according to different loads is seen Fig. 2. Increasing load causes more elongation of cantilever beam.

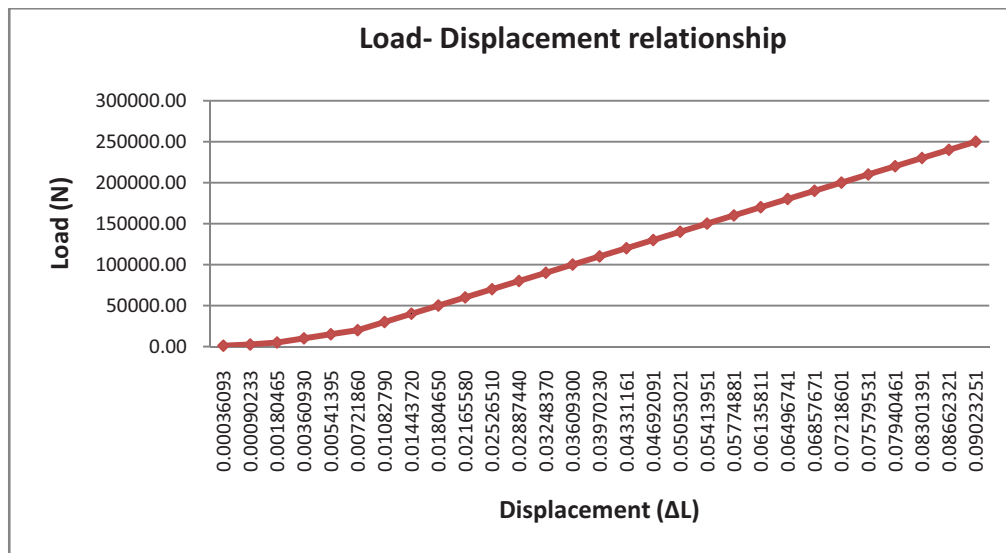


Figure 2. Variation of Load– Displacement

2.2 Regression Analysis

Regression analysis is performed for the modeling and analysis of several variables where there is a relationship between a dependent variable and one or more independent variables. The correlations between input parameters length, area, applied force, stress and output parameter strain are obtained by multiple linear regression. Through the backward elimination process, the final quadratic models of the response equation in terms of actual strain are presented as follows (Eq (5)):

$$\text{Strain}_{REGA} = (3E - 09x^2) - (2E - 07x) + (7E - 06), \quad R^2 = 0,9597 \quad (5)$$

The coefficient of determination (R^2) defines the correlation between experimental and predicted results. The differences between the real responses which are measured after the experiments and the estimated responses that are calculated with the above Eq. 5 is given. This REGA are presented [16] and this model has $R^2 = 0.9597$.

ANOVA helps in formally testing the significance of all main factors by comparing the mean square against an estimate of the experimental errors at specific confidence levels [16]. The

optimal combination of process parameters are predicted by both *S/N* and ANOVA analysis. The ANOVA results obtained for displacement (ΔL) according to parameters at the confidence level of 95% is performed.

The ANOVA table is a widely used model which tests the significance and interactions of the factors and the suitability of the model [17]. The F-Test is calculated in order to show by what amount the test results are affected by the control factors. In the ANOVA table, when “*Prob > F*” is less than 0.05, the control factors and their interactions are significant [18].

2.3 Finite Element Analysis (FEA)

Cantilever beam with different profiles is modelled in three dimensional (3D). In Fig. 3 (a), a cantilever beam model with I profile can be seen. After modelling, different axial loads are applied on cantilever beam that has different dimension. The first step of the of mesh optimization is model determination. This optimization continues until the FEA results and empirical solutions are close to each other. After the determination of mesh model, different axial loads are applied on cantilever beam that have different dimensions systematically. The cantilever beam meshed can be seen in Fig. 3 (b). Then maximum displacement values on the models are collected and classified according to model dimensions and applied loads. Then FEA analysis is performed and compared with the empirical solution model.

The elastic stress and strain fields of finite element method are systematically investigated using 3D finite element method. It is found that the maximum displacements of the finite cantilever beam are different and depend on type of material. The maximum stress and strain always occur on the fixed part of cantilever beam (Fig. 3 (c) and (d)). Displacement increase with increasing of length and decrease with increasing of cross sectional area. The contour of stress and strain are shown at FEA model.

A parametric FEA model is developed for determination of displacements in a cantilever beam with different profile types. First step is mesh optimisation for workpiece. Length, area and applied load, profile and material are chosen as analysis parameters. After that different tensile and compressive loads are applied the model, results are compared with empirical solution. Analysis results are collected and classified. FEA results have been compared with the others models. These models were Empirical model, Regression Analysis (REGA) model, Artificial Neural Network (ANN) model.

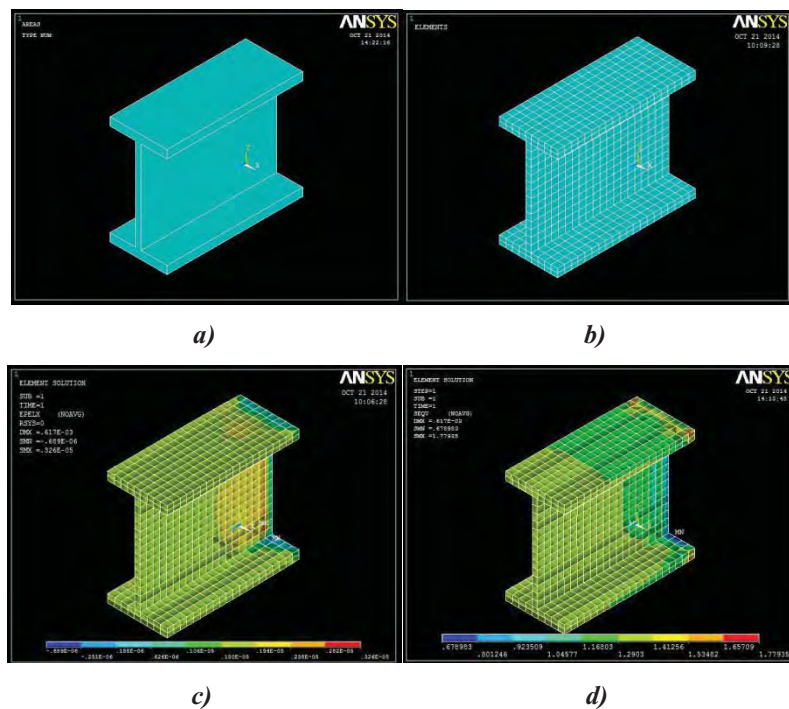


Figure 3. a) Modelling b) Meshing c) X component of elastic strain d) Von Mises Stress

2.3 Artificial Neural Network (ANN) Model

A neuron is the basic element of neural networks and depending on its duties its shape and size may vary. Analyzing a neuron in terms of its activities is important, since understanding the way it works also helps us to construct the ANNs. An ANN may be seen as a black box which contains hierarchical sets of neurons (e.g., processing elements) producing outputs for certain inputs [19].

Each processing element consists of data collection, processing the data and sending the results to the relevant consequent element. The whole process may be viewed in terms of the inputs, weights, the summation function, and the activation function (Fig. 4) [20].

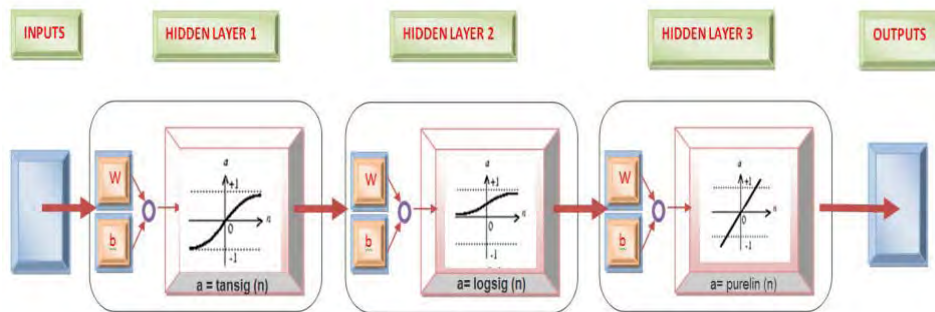


Figure 4. Basic artificial neural network model [19]

According to the Fig.4, we have the following:

- I. *The inputs*, are the activity of collecting data from the relevant sources.
- II. *The weights*, control the effects of the inputs on the neuron. In other words, an ANN saves its information over its links and each link has a weight. These weights are constantly varied while trying to optimize the relation in between the inputs and outputs.
- III. *Summation function*, is to calculate of the net input readings from the processing elements.
- IV. *Transfer (activation) function*, determines the output of the neuron by accepting the net input provided by the summation function. There are several transfer functions like summation function. Depending on the nature of the problem, the determination of transfer and summation function are made. A transfer function generally consists of algebraic equations of linear or nonlinear form [21]. The use of a nonlinear transfer function makes a network capable of storing nonlinear relationships between the input and the output. A commonly used function is sigmoid function because it is self-limiting and has a simple derivative. An advantage of this function is that the output cannot grow infinitely large or small [22].
- V. *Outputs*, accept the results of the transfer function and present them either to the relevant processing element or to the outside of the network.

The functioning of ANNs depends on their physical structure. An ANN may be regarded as a directed graph containing a summation function, a transfer function, its structure, and the learning rule used in it. The processing elements have links in between them forming a layer of networks. A neural network usually consists of an input layer, a number of hidden layers, and an output layer [20].

In this study, the training and test data have been prepared using empirical patterns. The data were obtained according to study parameters that has 107605 lines. Among them, 295 patterns each profile shape have been randomly selected and used as the test data. About 30% data were used as test data, 70 % data were used training data. Length, width, thickness, hole diameter and applied force were used as input-layer, LM algorithm and MLP (Multi Layer Perception) were used in the ANN model. Strain was used as output-layer of the ANNs. In the ANN model, tansig, logsig and purelin transfer functions (f) have been used and expressed as follows (Eqs 6-9):

$$NET_i = \sum w_{ij} \cdot x_j + w_{bi} \quad (6)$$

$$a = \text{tansig}(n) = \frac{2}{(1+e^{-2n})} - 1 \quad (7)$$

$$a = \text{logsig}(n) = \frac{1}{(1+e^{-n})} \quad (8)$$

$$a = \text{purelin}(n) \quad (9)$$

n : Number of processing elements in the previous layer

where NET is the weighted sum of the input. Input and output values (o) are normalized between 0 and 1.

Generally, there are three different learning strategies. Firstly, the trainer may tell the network what it should learn (supervised learning), secondly, the trainer may indicate whether or not the output is correct without telling what the network should learn (reinforcement learning), and finally, the network learns without any intervention of the trainer (unsupervised learning). The learning set consists of the inputs and the outputs used in training the network. The required outputs take place in this set in the case of supervised learning, while in other cases, they are not found in it [20]. In our case, we have used supervised learning approach. Since the number of neurons found in the input and output layers are known, the best performance of the network with the number of hidden layers is determined using trial error method. Using limited number of neurons with limited number of hidden layers causes lesser learning, while increasing these numbers too much, decreases the speed of learning, and in some cases prevents the learning entirely. Usually, an algorithm is used for the learning process, this algorithm determines the weights. There are various learning methods using these strategies [20].

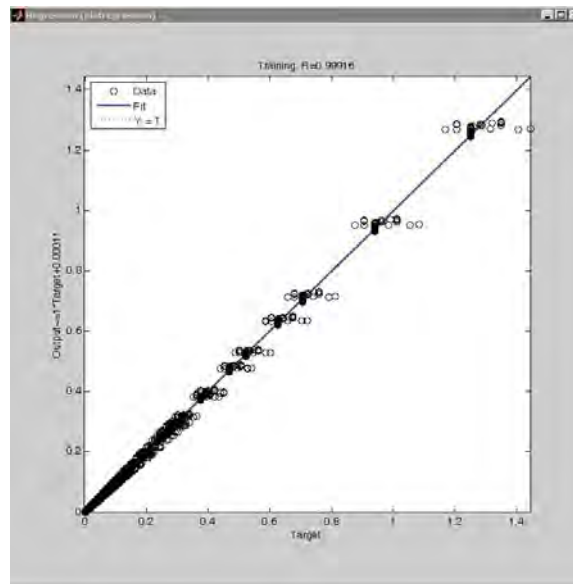


Figure 5. The ANN predictions against the empirical-based results.

The back propagation learning algorithm has been used with Scaled Conjugate Gradient (SCG) learning algorithm and Levenberg-Marquardt (LM) learning algorithm versions at the training and testing stages of the networks [23]. The computer program has been developed under MATLAB [24]. In the first step of the training, a determination of the learning algorithms is made. The number of hidden layers and the number of neurons for each hidden layer are determined. Then, the number of iterations are entered by the user, and the training starts. The training continues either to the end of the iterations or reaching the target level of errors. Fig.5 illustrates the ANN predictions against the empirical results.

3. TESTING THE ACCURACY OF REGRESSION ANALYSIS, ANN-BASED AND FINITE ELEMENT MODEL

In order to understand whether a multiple regression analysis or an ANN is making good predictions, the test data which has never been presented to the network is used and the results are checked at this stage. The statistical methods of $RMSE$, R^2 , and MEP values have been used for making comparisons [19,25,26]. The same data obtained from the regression analysis is used to determine the mentioned values.

These values are determined by the following Eqs (10-12):

$$RMSE = \left((1/p) \sum_j |t_j - o_j|^2 \right)^{1/2} \quad (10)$$

$$R^2 = 1 - \left(\frac{\sum_j (t_j - o_j)^2}{\sum_j (o_j)^2} \right) \quad (11)$$

$$MEP = \frac{\sum_j \left(\frac{t_j - o_j}{t_j} \times 100 \right)}{p} \quad (12)$$

where t is the target value, o the output, and p the number of samples.

Using the trial error method, the structure of the network (i.e. the number of neurons and hidden layers) is altered and the training operation is repeated. To be able to get accurate results we have used three hidden layers and for each hidden layer we have changed the number of neurons used at each hidden layer (e.g. from 5 to 150) to get the best network in terms of the statistical errors that it provides.

4. RESULTS AND DISCUSSION

In this study we have composed the empirical solution and network predicted strain results for the cantilever beam statistical error analysing methods. As presented in Table 3, the statistical error levels for both training and testing data sets are evaluated. As the table illustrates the network with four hidden layers of 4+9+9+11+17+1 neurons at each layer has provided the best results (Fig. 6).

Following, the ANN model as illustrated in Fig. 6 set up using 4 neurons for the input layer and with 9+9+11+17 processing elements at four hidden layers and finally 1 neuron are used at the output layer. The representation of knowledge is accomplished by the weights inbetween the layers.

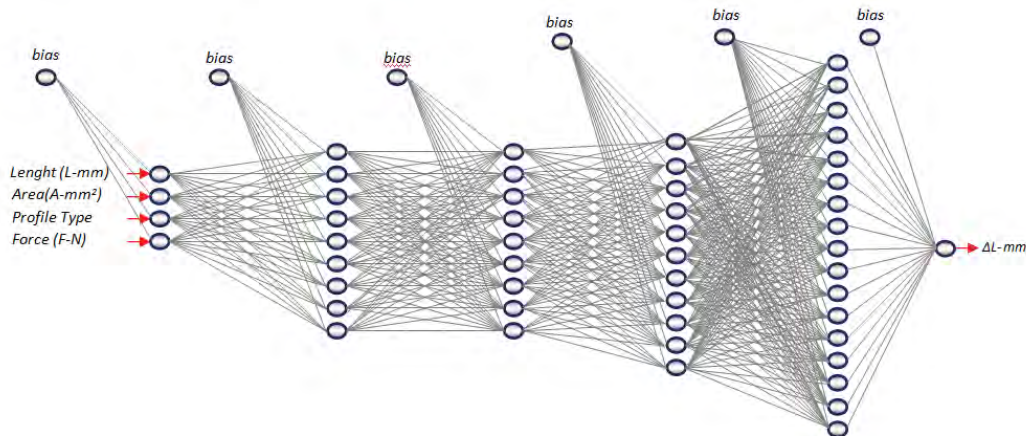


Figure 6. ANN architecture with 9+9+11+17 processing elements at four hidden layers

In terms of the statistical error analysis methods, using Levenberg-Marquardt (LM) learning algorithm technique for Strain, the mean error value for the training data set is 0.002921160% while for the testing data it becomes 0.003012259% (Table 3).

Table 3. Comparison of ANN model, REGA model, Analytical model and FEA with empirical solution using statistical error analysing methods for the cantilever beam model

| | | Absolute Fraction of Variance (R^2) | Root Mean Square Error (RMSE) | Mean Error Percentage (MEP %) |
|---|------|---|-------------------------------|-------------------------------|
| Displacement (ΔL) _{ANN} | Test | 0.999999182 | $3.6353 \cdot 10^{-17}$ | 0.003012259 |
| Displacement (ΔL) _{Regression} | - | 0.9597 | $2.98194 \cdot 10^{-13}$ | 0,942807101 |
| Displacement (ΔL) _{FEA} | - | 0.999739705 | $2.4984 \cdot 10^{-10}$ | 0.080527482 |

Analysis results have been collected and classified according to the length of the cantilever beam (L), area of profile, material and the applied axial load (F). FEA strain results have been compared with the others models. These models were empirical model, regression analysis (REGA), Artificial Neural Network (ANN) model.

This optimization has continued until the FEA results and empirical solutions close to each other. Maximum stresses are collected and classified according to model dimensions and applied load. Then FEA have been performed and compared with the empirical results.

Variation in strain (ϵ) according to Modulus of Elasticity (E) can be seen in Fig. 7. Increasing of E leads to decreasing of ϵ with respect to graph.

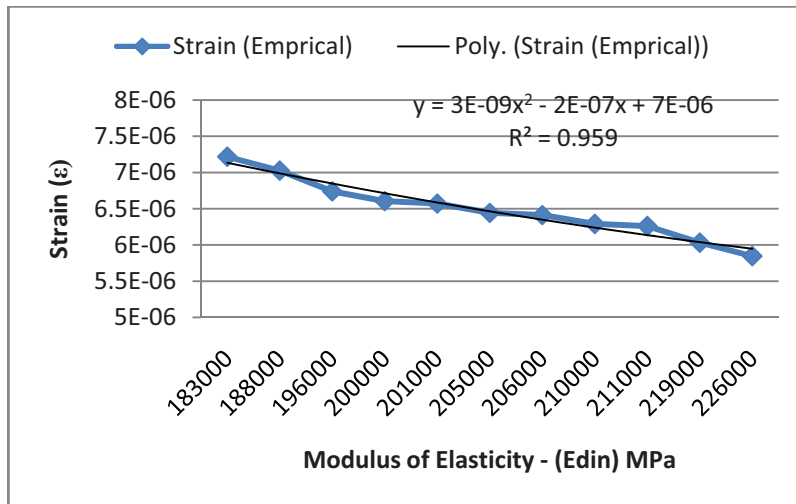


Figure 7. Variation of Modulus of Elasticity - Strain

According to material type (for different E , relation between displacement (ΔL) and strain (ϵ) are shown in Fig. 8. While highest strain rate occurs at $E=183000$ MPa, and the lowest strain rate comes up at $E=226000$ MPa.

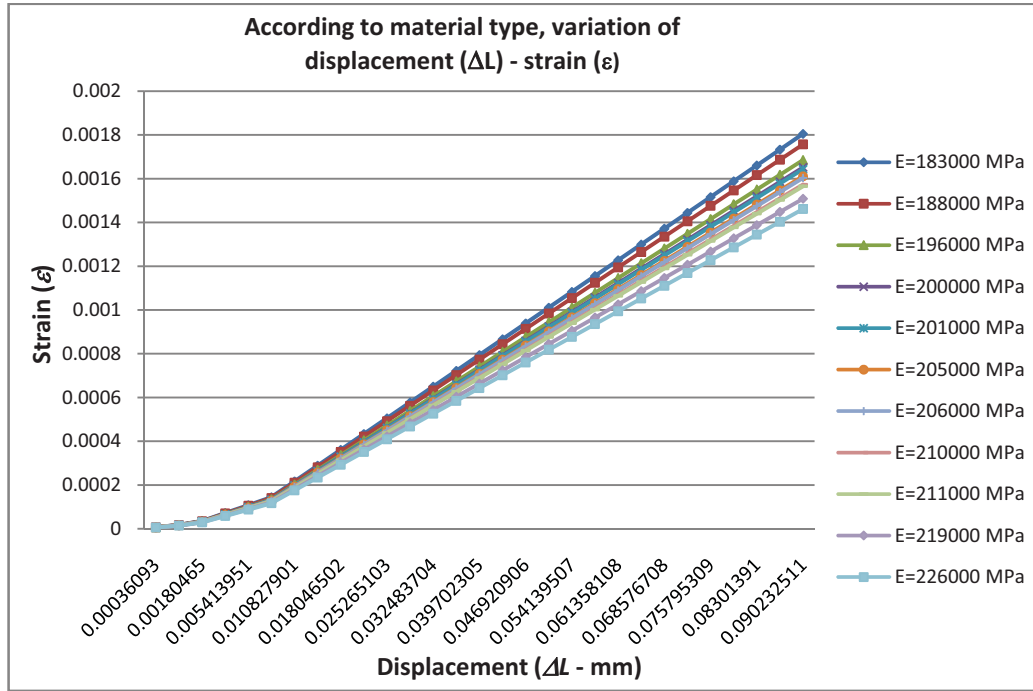


Figure 8. According to material type, variation of displacement (ΔL) - Strain (ϵ) chart (I profile type, $L=50\text{mm}$, $A=757\text{mm}^2$)

5. CONCLUSIONS

In this study, the application of regression analysis, ANN and FEA on the empirical displacement (ΔL) values were compared and discussed. This study was revealed some inaccuracy in empirical model and FEA. The developed models which are limited with their boundary conditions are compared in terms of the prediction accuracy to displacement (ΔL). For a long time, modeling techniques have been developed for prediction of the displacement (ΔL). From the results of this study, the following conclusions are drawn. The best results were obtained using ANN method. It is concluded that ANNs can be used as an alternative method to determination of displacement for cantilever beam.

6. REFERENCES

- [1] Sha W., and Edwards K. L., "The use of artificial neural networks in materials science based research," in *Materials & Design*, vol. 28, Issue 6, 2007, pp. 1747-1752.
- [2] Toktaş İ., "Mekanik Sistemler için bir Kavramsal Tasarım Modelinin Geliştirilmesinde Yapay Sinir Ağlarının Kullanılması", Ph.D. Thesis, Gazi University, Institute of Science and Technology, 2003.
- [3] Prabhakar S., and Henderson M. R., "Automatic Form-Feature Recognition Using Neural-Network-Based Techniques on Boundary Representations of Solid Models," in *Computer Aided Design*, vol. 24 (7), 1992, pp. 38-393.
- [4] Henderson M. R., "Manufacturing Feature Identification," in *Artificial Neural Networks for Intelligent Manufacturing*, Ed. Daglı, C.H. Chapman & Hall, Great Britain, Cornwall, 1994, pp. 229-264.
- [5] Nezis K., and Vosniakos G., "Recognizing 2½D Shape Features Using a Neural Network and Heuristics," in *Computer Aided Design*, vol. 29 (7), 1997, pp. 523-539.
- [6] Shtub A., and Versano R., "Estimating the cost of steel pipe bending, a comparison between neural networks and regression analysis," in *International Journal of Production Economics*, vol. 62, Issue 3, 1999, pp. 201-207.
- [7] Waszczyszyn Z., and Ziemiański L., "Neural networks in mechanics of structures and materials – new results and prospects of applications," in *Computers and Structures*, vol. 79, 2001, pp. 2261-2276.

- [8] Suresh S., Omkar S. N., Ganguli R., and Mani V., "Identification of crack location and depth in a cantilever beam using a modular neural network approach," in *Smart Materials and Structures*, vol. 13 number 4, 2004, pp. 907.
- [9] Gholizadeh S., Pirmozb A., and Attarnejad R., "Assessment of load carrying capacity of castellated steel beams by neural Networks," in *Journal of Constructional Steel Research*, vol. 67, 2011, pp. 770-77.
- [10] Rybicki E. F., and Kanninen M. F., "A finite element calculation of stress intensity factors by a modified crack closure integral," in *Engineering Fracture Mechanics*, vol. 9, Issue 4, 1977, pp. 931-938.
- [11] Nguyen D.K., "Large displacement behaviour of tapered cantilever Euler–Bernoulli beams made of functionally graded material," in *Applied Mathematics and Computation*, vol. 237, 2014, pp. 340-355.
- [12] Yang Q., Zheng B.L., Zhang K., and Li J., "Elastic solutions of a functionally graded cantilever beam with different modulus in tension and compression under bending loads," in *Applied Mathematical Modelling*, vol. 38, Issue 4, 2014, pp. 1403-1416.
- [13] Macneal R. H., and Harder R. L., "A proposed standard set of problems to test finite element accuracy," in *Finite Elements in Analysis and Design*, vol. 1, Issue 1, 1985, pp. 3-20.
- [14] Ghafoori E., and Motavalli M., "Analytical calculation of stress intensity factor of cracked steel I-beams with experimental analysis and 3D digital image correlation measurements," in *Engineering Fracture Mechanics*, vol. 78, 2011, pp. 3226-3242.
- [15] Mukavemet ve Malzeme (2009) Citation Reference [Online]. Available: [http:// www.tasarimveimalat.com/mukavemet.pdf](http://www.tasarimveimalat.com/mukavemet.pdf)
- [16] Ross P. J., "Taguchi Techniques for quality engineering", Mc Graw Hill International Edition, Singapur, 1996.
- [17] Bas D., and Boyaci I., "Modeling and optimisation. II. Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction," in *J. Food Eng.*, vol. 78, 2007, pp. 846-854.
- [18] Ko-Ta C., "Modeling and analysis of the effects of machining parameters on the performance characteristics in the EDM process of Al₂O₃+TiC mixed ceramic," in *The International Journal of Advanced Manufacturing Technology*, vol. 37, 2008, pp. 523-533.
- [19] Toktaş İ., and Başak H., "Chain Gears Design Using Artificial Neural Networks," in *Computer Applications in Engineering Education*, vol. 20, issue 1, 2012, pp. 38-44.
- [20] Oztemel E., "Integrating expert systems and neural networks for intelligent on-line statistical process control", Ph.D. Thesis, School of Electrical, Electronic and Systems Engineering, University of Wales, Cardiff, Wales, UK, December 1992.
- [21] Hollingworth N., and Hills D. A., "Theoretical efficiency of a cranked link chain drive," *Proc Inst Mech Eng Part C*, 200, 1986, pp. 375_377.
- [22] Massie D. D., "Neural network fundamentals for scientists and engineers," in *Proceedings of the International Congress on Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems and Processes (ECOS '01)*, Istanbul, Turkey, 2001, pp. 123-128.
- [23] Hagan M.T. and Demuth H. B., "Neural Network Design", PWS, Boston, Mass, USA, 1996, pp. 1-29.
- [24] MATLAB 6.5 (Release 13), "The language of technical computing", The MathWorks, Inc., Natick, MA, 2002.
- [25] Ozkan M.T., "Experimental and artificial neural network study of heat formation values of drilling&boring operations on Al 7075 T6 workpiece," in *Indian Journal of Engineering & Materials Science*, vol. 20 (4), 2013, pp. 259-268.
- [26] Ozkan M.T., Eldem C., and Sahin I., "Determination of Notch Factor for Shafts Under Torsional Stress by the Help of Artificial Neural Networks," in *Materiali in tehnologije / Materials and Technology, MTAEC9*, vol. 48 (1), 2014, pp. 81-90.

A Novel Method for Islanding Detection of Distributed Generation Units

Chinedu Frank Okwose ¹, Reza Sirjani ²

^{1,2} Faculty of Engineering, Cyprus International University,
Nicosia, Northern Cyprus, Mersin 10, TURKEY

¹ *franqudff@yahoo.com*, ² *rsirjani@ciu.edu.tr*

ABSTRACT

Islanding is a situation in which the distributed generation (DG) unit becomes isolated from the main power system due to inadvertent opening of circuit breakers. Islanding detection is an ongoing challenge because existing methods are not entirely satisfactory. This paper presents a novel islanding detection method in which active and reactive power fluctuations are simultaneously measured in time intervals. To demonstrate the accuracy and speed of proposed method, a wind turbine is modeled in Power World Simulator Software and islanding detection time obtaining by proposed method is compared with the results of other methods.

Keywords –Distributed generation, Islanding detection, Active power, Reactive power

1. INTRODUCTION

Over the past decade renewable energy has been intensively developed and it generates lower pollution than do fossil fuels and nuclear generation systems. The new paradigm of distributed generation (DG) thus increases in technical importance and increases profits globally. In principle, DG is a small-scale generation unit installed to the load and connected to the grid, for selling or buying of energy [1]. DG units such as photovoltaic generation, wind power generation, small gas turbine generation, fuel-cell generation, cogeneration and etc. operate in parallel with power distribution system. This parallel operation of a dispersed generation unit with the utility distribution system may bring about some critical problems, that is, difficulties and complications for the reliable and safe operation of both systems [1, 2]. One of those problems is islanding. The islanding condition occurs when a portion of the utility system that contains both load and distributed resources remains energized while it is isolated from the remainder of the utility system. IEEE standard recommends disconnecting all distributed generators immediately after the formation of island [3]. Islanding can be intentional or unintentional. Intentional islanding may be due to preplanned event such as maintenance and in case of an unintentional island, the DG should be disconnected within 2 seconds after the grid failure [1, 4].

Islanding is generally defined as a situation where the power from utility is off, but one (or more) sections of the system still continues to have power flow through it as it still being energized by DG. This sectionalized area is called an island. Unintentional islanding refers to a formation of island due to faults on the utility side that result in the opening of the circuit breaker in the upper stream of the grid [1]. Fast and accurate detection of islanding is one of the major challenges in today's power system with many distribution systems already having significant penetration of DG as there are few issues yet to be resolved with islanding. Islanding detection is also important as islanding operation of distributed system is seen a viable option in the future to improve the reliability and quality of the supply. Islanding detection schemes are commonly evaluated based on the Non-Detection Zone (NDZ). The NDZ corresponds to the range of active and reactive load-generation mismatches within the island in which the islanding detection approach fails to identify the islanding state.

The islanding detection schemes can be grouped into two categories: remote and local as shown in Fig. 1.

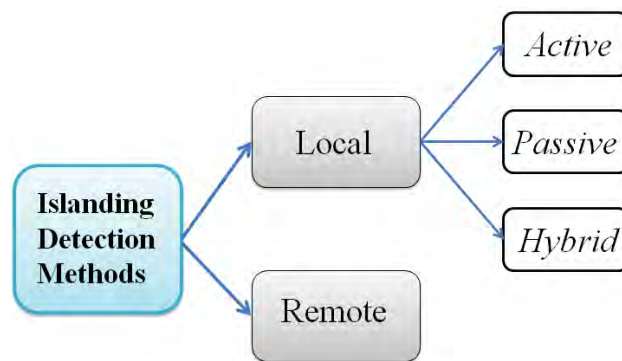


Figure 1. Classification of islanding detection schemes

Remote techniques are based on communication between the electric utility and the DG units. Despite the fact that remote techniques are reliable and effective, they suffer high implementation cost. In multi-inverter systems it is effective but expensive to implement (especially in small systems) and has a complicated communication technique [2].

Local islanding schemes can further be divided into active, passive and hybrid.

Active methods: They rely on injecting perturbations in the distribution system to facilitate significant changes in the power system parameters and hence allow easy detection of the island. Active techniques have small NDZ, but their operation results in degrading the power quality because they introduce perturbations in the voltage and/or current at predefined intervals which defeats the objective of having digital-grade power quality attribute as aimed in smart grid [4].

Passive methods: They are based on local measurements of power system parameters at the point of common coupling of the DG. Passive methods detect islanding conditions by measuring changes in the electrical quantities at the DG output. Unlike active methods, passive methods are inexpensive, easy to implement due to reduced complexity and maintain the quality of power [5-7].

Hybrid methods: They are combinations of both active and passive schemes. They introduce perturbations through active methods only after the detection of the island by passive scheme and thus, reducing the amount of perturbations injected into the system. However, hybrid methods need longer time to detect the island compared to active and passive methods [4].

Existing active techniques are generally very effective in detecting islanding phenomenon, but have the significant disadvantage of requiring a direct influence on power system such as overcurrent, voltage quality and etc. Because the exported real and reactive power must be large enough to perfectly operate the over/under frequency relay for detecting islanding operation. Passive techniques can avoid this problem and are generally the least expensive to install, but they cannot be guaranteed to operate under all islanding condition. In [8], the authors proposed a method that uses voltage magnitude variation to detect islanding phenomenon by monitoring the fluctuations in the reactive power and voltage magnitude at the inter-tie point caused by varying the internal induced voltage of DG unit. In [9] the authors proposed a new methodology based on the active and reactive powers changes to detect the islanding state. In this method if the active power is less than threshold, system will continue to work. But if the real power value is more than threshold, reactive power value is considered and if its value exceeds threshold as well, islanding is detected and disconnect command will be transmitted.

In this paper a novel method for detecting the islanding condition in DG units is proposed. This method simultaneously measures active and reactive power variations. A new islanding detection factor for computing complex power variations is defined and calculated in each time intervals. The proposed method is implemented on two case studies and the results are compared with other methods.

2. METHODOLOGY

By monitoring the fluctuations in both active and reactive powers at the inter-tie point, islanding phenomenon can be detected. In general, synchronous generator is used for DG unit as shown in Fig. 2.

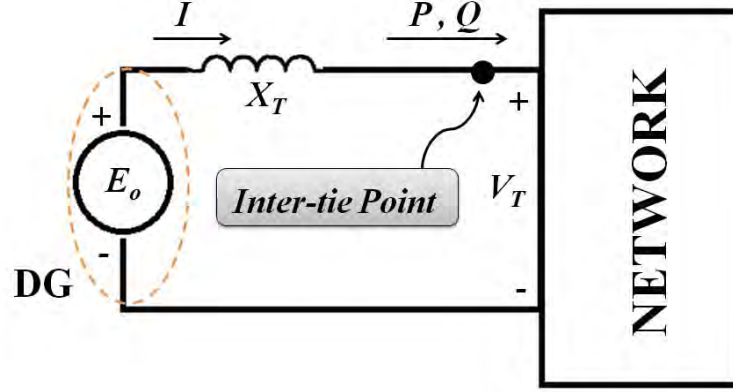


Figure 2. Equivalent circuit of network with DG unit

The active and reactive power P , Q supplied to the network from the DG unit can be calculated as follows [8] :

$$P = \frac{|E_o||V_T|}{X_T} \sin\delta \quad (1)$$

$$Q = \frac{|V_T|}{X_T} (|E_o| \cos\delta - |V_T|) \quad (2)$$

Where: E_o is internal induced voltage in DG unit, V_T is system voltage, X_T is synchronous reactance per phase and δ is difference phase angle between E_o and V_T .

On the other hand, we know the apparent power (S) is a complex variable which is expressed by P and Q as follow:

$$S = P + jQ \quad (3)$$

Therefore the fluctuations in both active and reactive powers can be expressed by load variations at inter-tie point:

$$\Delta S = \Delta P + j\Delta Q \quad (4)$$

Now, we define a new islanding detection factor to compute apparent power variations at specific time, based on the values of P and Q at the first time interval:

$$S_{istd.} = \frac{\sqrt{\Delta P_K^2 + \Delta Q_K^2}}{\sqrt{P_1^2 + Q_1^2}} \quad (5)$$

Where, P_1 and Q_1 are active and reactive powers in the first time interval (t_1) at the inter-tie point. ΔP_K and ΔQ_K are active and reactive power variations at time t_K .

The proposed procedure for islanding detection based on fluctuations in both active and reactive powers are as follow:

- i. Measure P and Q in the first time interval at the inter-tie point. (P_1 and Q_1)
- ii. Measure P and Q in next time interval at inter-tie point. (P_K and Q_K)
- iii. Compare the values of P and Q with the previous time interval and calculate ΔP_K and ΔQ_K .
- iv. Calculate ΔS_K or load variations at Inter-tie point: $\Delta S_K = \Delta P_K + j \Delta Q_K$
- v. Calculate islanding detection factor (S_{isld}) using Eq. (5).
- vi. If $S_{isld} > 1$, the islanding has been occurred and the breaker should trip.
- vii. If $S_{isld} < 1$, go to step (ii).

Flowchart of proposed method for islanding detection is shown in Fig. 3.

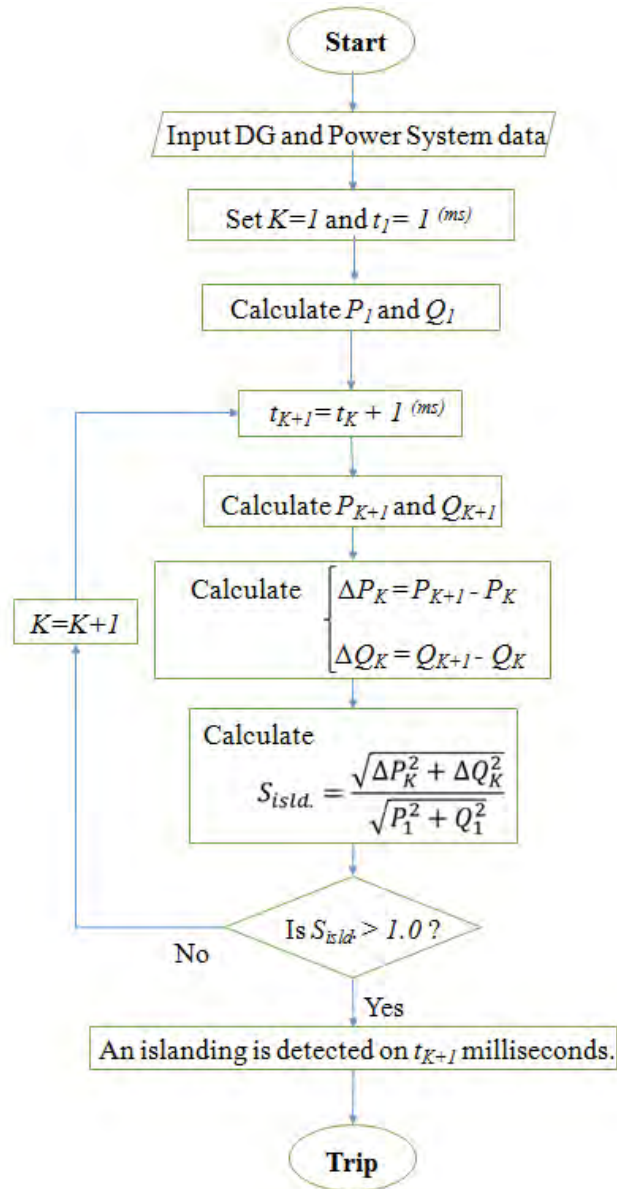


Figure 3. Flowchart of proposed method for islanding detection

3. CASE STUDY AND RESULTS

The single line diagram of system considered in this paper is shown in the Fig. 4 [10]. A DG unit is considered as generator in order to represent a wind farm with the initial current into the infinite bus set to 1.0 (unity power factor). The generator reactance is $X_{eq}=0.8$ per unit using a 100 MVA system base.

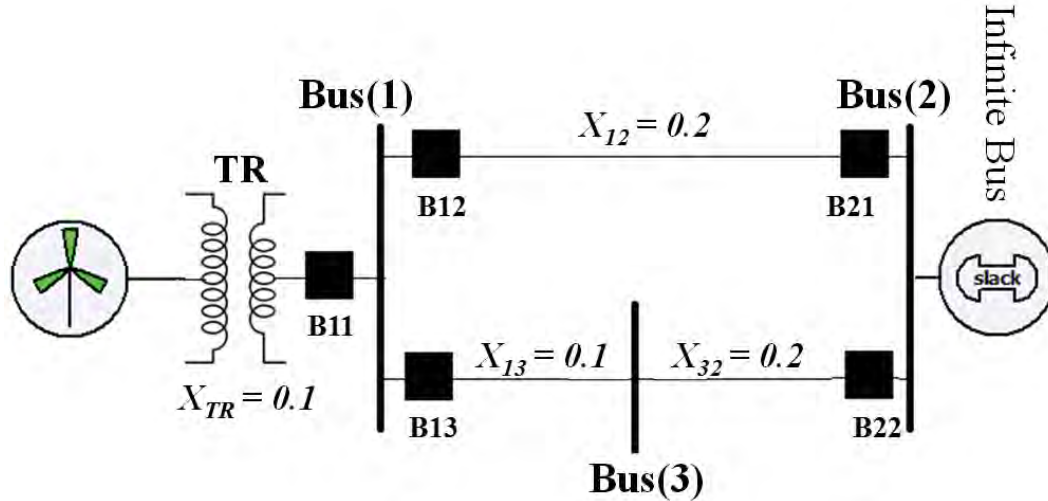


Figure 4. Single line diagram of test system

The wind turbine DG of 186MW is modeled in Power World Simulator Software [10] and simulation results are given in Fig 5. In this case study, at the first time interval (One millisecond), DG supplies active and reactive power of $P_1=57.489\text{MW}$ and $Q_1=34.982\text{MVar}$, respectively, and the grid supplies the remaining power.

The islanding detection factors ($S_{isld.}$) are calculated for each time interval and all of them are less the 1, before 168 milliseconds. At the time 168 milliseconds, active and reactive power variations reached to $\Delta P_{168}=12.327\text{MW}$ and $\Delta Q_{168}=77.219\text{MVar}$, respectively and the S_{isld} value is 1.162 which is greater than 1. It means an islanding has been occurred at time 168 milliseconds and the breaker should trip.

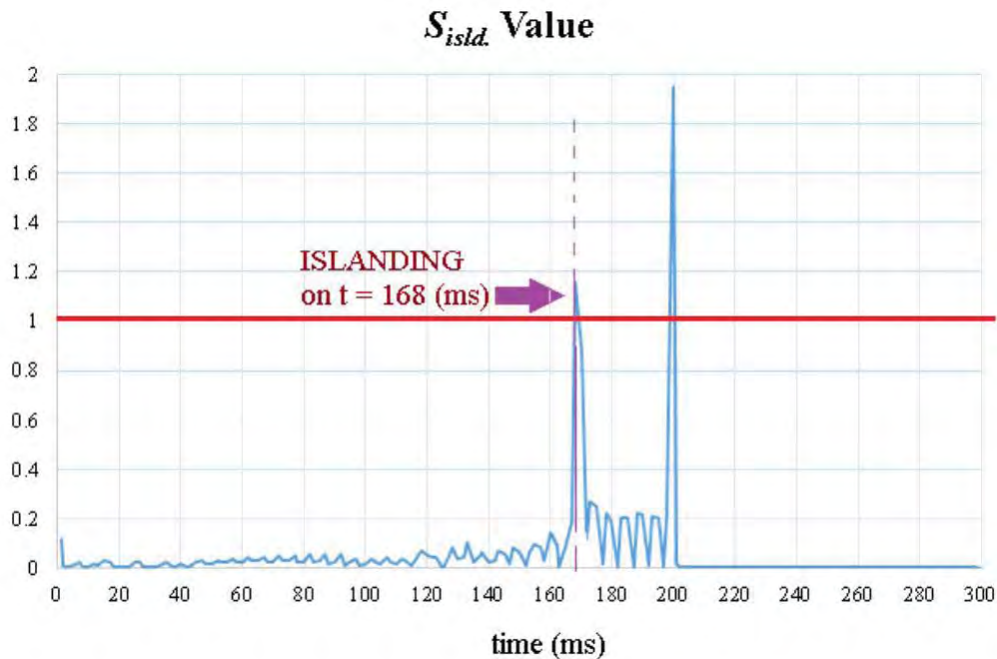


Figure 5. Values of islanding detection factor for wind turbine DG of 186MW

The same procedures have been done on wind turbine DG of 200 MW in Power World Simulator Software and simulation results are given in Fig 6. Here, at the time 18 milliseconds, the S_{isld} value reached to 1.196 which is greater than 1. It means an islanding has been occurred at time 18 milliseconds and the breaker should trip.

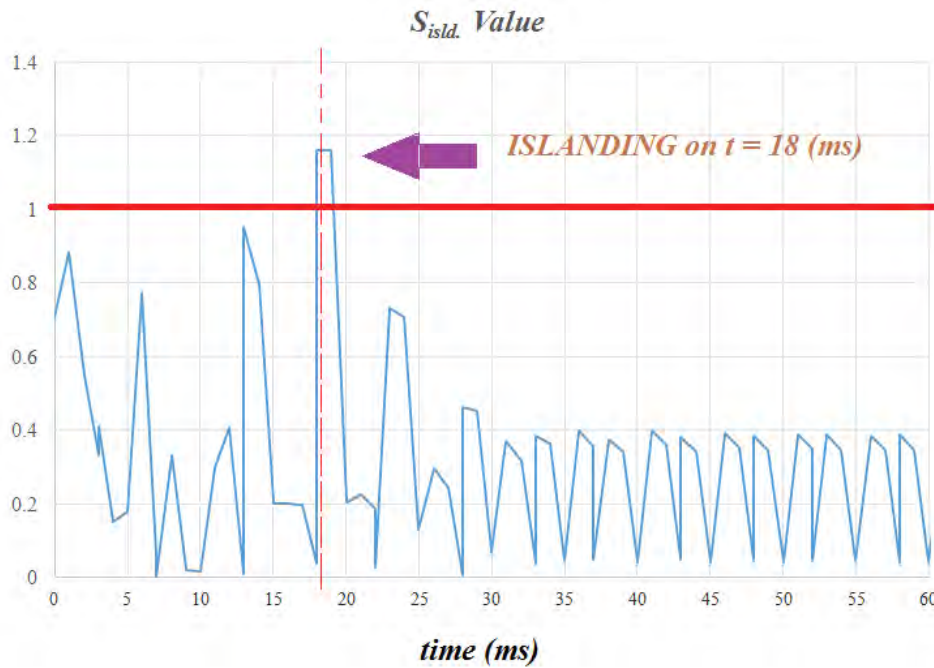


Figure 6. Values of islanding detection factor for wind turbine DG of 200MW

In order to demonstrate the accuracy and speed of the proposed method, two other islanding detection methods have been implemented on the aforementioned case studies and the results have been compared. These methods are as follow:

1. **Method I** : Monitoring the fluctuations in the reactive power and voltage magnitude [8]
2. **Method II** : Combined changes of active and reactive powers of distributed generations [9]

The islanding detection time using different methods are presented and compared in Table 1. It is observable that the proposed method in both case studies is relatively faster than *Method I* and *Method II*.

Table 1. Comparison of islanding detection time using different methods

| Case Study | Islanding Detection Time | | |
|------------|--------------------------|---------------------------|-----------------|
| | <i>Method I</i> [8] | <i>Method II</i> [9] | Proposed Method |
| DG : 186MW | 171 ms | 200 ms | 168 ms |
| DG : 200MW | 20 ms | 19 ms | 18 ms |

4. CONCLUSIONS

In this paper a new method to detect islanding condition of DG has been proposed. This method measures the fluctuations in both active and reactive powers in same time. A new islanding detection factor to compute apparent power variations at specific time has been defined and calculated in each time interval. A wind turbine DG has been modeled in Power World Simulator Software and islanding detection time has been determined and compared with the results of different methods. The results demonstrated the significant speed of proposed method.

5. REFERENCES

- [1] B. Yu , M. Matsui and G.A. Yu, "A review of current anti-islanding methods for photovoltaic power system". *Solar Energy*, vol. 84, 2010 , pp.745–54.
- [2] W.G. Morsi, C.P. Diduch, C. Liuchen, and M.E. El-Hawary, "Wavelet-based reactive power and energy measurement in the presence of power quality disturbances," *IEEE Transactions Power Delivery*, vol. 26, no. 3, , 2011, pp. 1263-1271.
- [3] IEEE 1547-2003 Standard for Interconnecting Distributed Resources with Electric Power Systems, 2003.
- [4] M. Vaziri, S. Vadhva, T. Oneal, and M. Johnson, "Smart grid, distributed generation, and standards," *IEEE Power and Energy Society General Meeting*, 2011, pp. 1-8.
- [5] P. Mahat, Z. Chen, and B. Bak-Jensen, "Review of islanding detection methods for distributed generation," in *Proc. of the 3rd International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, 2008, pp. 2743-2748.
- [6] S.P. Chowdhury, S. Chowdhury, C.F. Ten and P.A. Crossley, "Islanding protection of distribution systems with distributed generators — A comprehensive survey report," in *Proc. of the IEEE Power and Energy Society General Meeting Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1-8.
- [7] F. De Mango, M. Liserre, A.D. Aquila and A. Pigazo, "Overview of antiislanding algorithms for PV systems. Part I: Passive methods," in *Proc. of the 12th International Conference on Power Electronics and Motion Control*, 2006, pp. 1878-1883.
- [8] J.E. Kim and J.S. Hwung , "Islanding detection method of distributed generation units connected to power distribution system" , *Power System Technology, International Conference on Power System Technology(PowerCon 2000)* ,2000 , pp. 643-647.
- [9] H.N. Aghdam , N. Ghadimi, P. Farhadi, F. Hashemi and R. Ghadimi, "Detecting the anti-islanding protection based on combined changes of active and reactive output powers of distributed generations",*3rd International Conference on Computer Research and Development (ICCRD)*,2011,pp. 285-289.
- [10] J.D. Glover, M.S. Sarma and T. Overbye, *Power System Analysis and Design*, Cengage Learning, Fifth Edition, 2011.