***Dosya Organiasyonunda Hashing Uygulaması***

1.Bedirhan ŞİŞMAN 2.Miraç Sıla GENÇ

bedirhan.sisman@gmail.com [gncsla@gmail.com](mailto:gncsla@gmail.com)

Bilgisayar Mühendisliği Bölümü

Kocaeli Üniversitesi

**ÖZET:**

Bu projeyi geliştirirken grafik kütüphanesi zorunlu olmadığı için kullanıcı dostu ara yüzü kodlarken C dilini kullanmaya karar verdik. Kodumuzun oluşturduğu 500 satırlık lineer liste içinden arama yapılabilmesi için boşlukta okumayı bırakan scanf kullanmanın kolay olacağını düşündük. İlk zamanlar programda derleme ile ilgili bir sürü hata ile karşılaştık. [1] Dosya için kullanılan C kütüphane kodlarını öğrenerek hatalarımızı düzelttik. Sonrasında algoritmaların karşılaştırılması için 500 elemanlı bir rasgele (random) kayıt dosyası oluşturduk ve bu dosyadan bazı C kodlarını kullanarak okuma gerçekleştirdik. Oluşturulan dosyadan 6 farklı biçimde tekrar dosya oluşturmakta çok zorlandık. Çünkü aynı zamanda hem random dosyadan bilgi alacak hem de bu bilgileri diğer 6 dosyada kullanacaktık. Çakışma oluştu ve çakışmanın giderilmesi için Hashing fonksiyonunu kullandık. [2] Ancak bağıl dosya organizasyonunu da küçük bir araştırma yaparak kodu yazmayı başardık. Bu adımlardan sonra en hızlı aramayı yapan yöntemi ekrana yazdırmak için döngülerin içine sayaç değişkeni atayarak gayet kolay bir şekilde kodumuzu sonlandırdık.

**1.GİRİŞ:**

Projemiz 500 elemanlı kayıt dosyası için yapılan en hızlı aramayı ( Bölen Kalan ve Lineer Yoklama, Bölen Kalan ve Ayrık Taşma, Katlama ve Lineer Yoklama, Katlama ve Ayrık Taşma, Kare ortası ve Lineer Yoklama, Kare ortası ve Ayrık Taşma ) ekrana yazdıran algoritmayı içerir. Daha sonra dosya içindeki verilerde arama yaparak yöntemleri karşılaştırmaktadır.

**2. TEMEL BİLGİLER:**

Kodumuzu oluştururken C dili kullanacağımız için Dev-C++ IDE'sini kullanmayı tercih ettik.[3]

Dev C**++** [GNU GPL](https://tr.wikipedia.org/wiki/GNU_GPL) altında dağıtılan [C (programlama dili)](https://tr.wikipedia.org/wiki/C_(programlama_dili)) ve [C++](https://tr.wikipedia.org/wiki/C%2B%2B) dilleri ile [GCC](https://tr.wikipedia.org/wiki/GCC)'nin Windows sürümü olan [MinGw](https://tr.wikipedia.org/w/index.php?title=MinGW&action=edit&redlink=1) [derleyici](https://tr.wikipedia.org/wiki/Derleyici) setini destekleyen [freeware](https://tr.wikipedia.org/wiki/Freeware) bir [tümleşik geliştirme ortamıdır](https://tr.wikipedia.org/wiki/Tümleşik_geliştirme_ortamı).

Dev C++' nin tüm arabirimi [Delphi](https://tr.wikipedia.org/wiki/Delphi) ile geliştirilmiş ve ilk çıktığı zaman oldukça popüler bir yazılım olmuştur. Proje [SourceForge](https://tr.wikipedia.org/wiki/SourceForge) tarafından barındırılmaktadır. İçinde dahili olarak bulunan DevPaks paket dağıtım aracıyla ek kütüphaneler, şablonlar ve araçlar programlama ortamına aktarılabilir. Güncel sürümü 5.8.1'dir. [1]

**3. GELİŞTİRİLEN MİMARİ:**

Kod geliştirilirken 7 tane dosya ile işlem yapılmıştır. İlk olarak random linner olarak oluşturulan kayıt dosyası oluşturulmaktadır. Diğer dosyalara bundan türetilen lineer ve ayrık taşma yoluyla elde edilmiştir.

**3.1 Hahsing Fonksiyonu**

Hash fonksiyonun görevi kendisine verilen bir değerden, benzersiz bir tamsayı yani key üretmesidir. Fakat uygulama alanında her zaman benzersiz bir sayı üretecek uygun fonksiyonu bulmak çok zor hatta bazı durumlarda imkansızdır. Eğer farklı iki değerden aynı sayı üretilirse bu duruma çatışma denir. Hashing fonksiyonundan beklenenler : [2]

* Herhangi bir uzunlukda değer alabilmelidir.
* Çıktı olarak belirlenen uzunlukta anahtar (key) üretebilmelidir.
* Tek yönlü olmalıdır. Fonksiyon tarafından üretilen anahtardan fonksiyona verilen değer elde edilmemelidir.
* Çatışmalara olanak vermemelidir.

**3.2 Hash Tablosu**

Verinin ve hasing fonksiyonuyla elde edilen değerlerin birlikte tutulduğu veri yapısıdır. [1]

**3.3 Bölen Kalan ve Yoklama**

Bu yönteme göre kullanıcının girmiş olduğu textteki sayının maximum kayıt sayısına göre modunu alıp çıkan modun değerine göre o satıra ekleme işlemi yapar.[2]

**3.4 Katlama Yoklama**

Bu yöntemde ise göre kullanıcının girmiş olduğu textteki 9 basamaklı sayının basamakları 3'er 3'er ayrılır daha sonra bu sayılar toplanır ve maximum kayıt sayısına göre modu alınıp çıkan değere o dosyada o satıra ekleme yapılır.[3]

Biz projemizde random olarak atanan öğrenci numaralarını 3 lü gruplamayı başaramadığımız için her bir rakamı kendi başına bir grup olarak düşünüp o şekilde bu yöntemi uygulamayı başarabildik.

**3.5 Kare Ortası Yoklama**

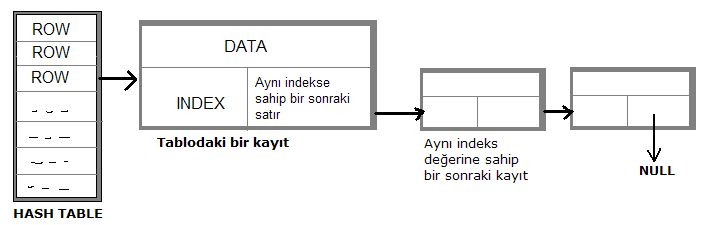
Bu yönteme göre yapılan işlem kullanıcının girmiş olduğu textte ki sayının karesi alınır daha sonra bu sayının basamak sayısı ve ortasındaki sayı bulunur ve o sayının maximum kayıt sayısına göre modu alınıp çıkan değere göre dosyada o satıra kayıt eklenir.[3]

**3.4 Lineer ve Ayrık Taşma Yöntemleri**

Yukarıda anlatılan dosyalama yöntemleri lineer ve ayrık taşmalı olarak her biri 2 farklı şekilde yapılmıştır.[3]

Lineer olarak yapılan yöntemde random olarak belirlenmiş kayıt dosyasındaki verilerin çakışması durumun da dosyada ki bir sonraki boşluğa atanır.

Ayrık taşma yönteminde dosya oluşturulurken en başta boşluklar bırakılır. Bu boşluk sayısı kodlayıcı tarafından belirlenir. Kayıt dosyasında çakışma olması durumunda çakışan veriler sırayla dosya oluşturulurken bırakılan boşluğa kaydedilirler.[4]

**

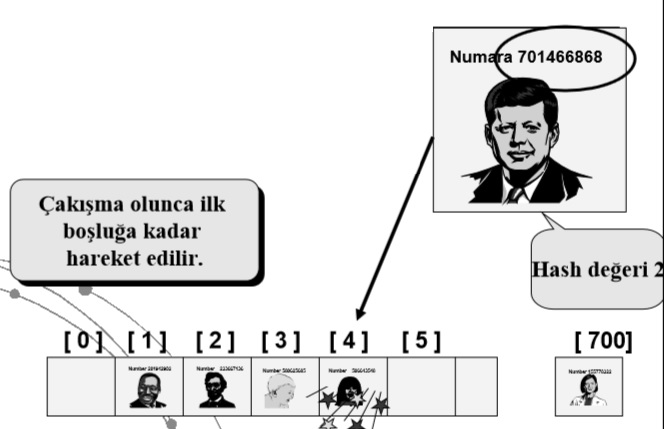
*Şekil1:Lineer Yoklama Yöntemi*

*Şekil2:Lineer Yoklama Yöntemi*

Gösterilmiş olan şekiller lineer yoklama yöntemiyle yapılan dosyalama işleminin temsili anlatımını temsili Olarak ifade etmek için kullanılmıştır.

**3.5 Yazılım Geliştirme**

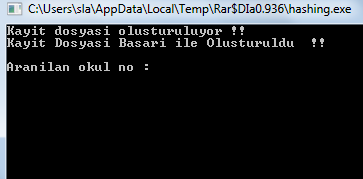
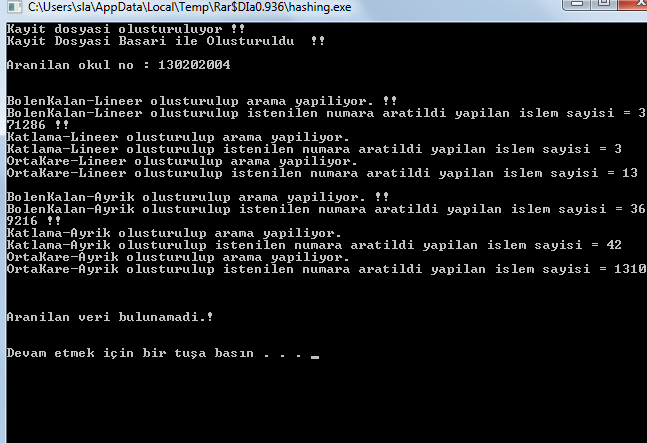
Yazılım geliştirme aşaması her 2 grup üyesi için eşit zaman harcanmıştır.

Projenin öncelikle konusunu kavramak için internet ortamından ve ders kaynakarından yararlanılmıştır.

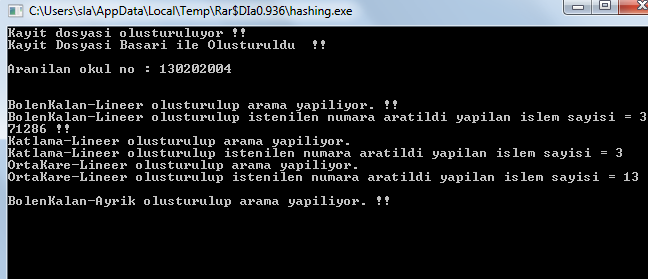
Konu kavrama aşamasından sonra kullanılacak olan programlama dili seçimi yapılmış daha sonra kodlamaya geçilmiştir. Kodlama gerçeklenirken grup üyeleri arasında ayrı ayrı çalışmalar yerine haftada dört yada beş gün ikişer saatlik çalışmalar sonunda genel bir hal elde edilmiştir.

# 4. Kullanıcı Kataloğu

Bu kısımda projemizin çalışır haldeki görüntüleriyle temel olarak nasıl çalıştığı ve ne şekilde kullanılacağı gösterilecektir.



Yukarıda yer alan ekran görüntüsünde ilk olarak çalıştırdığımız hali karşımıza gelmektedir. random olarak bir kayıt dosyası başarılı bir biçimde oluşturulmuştur.



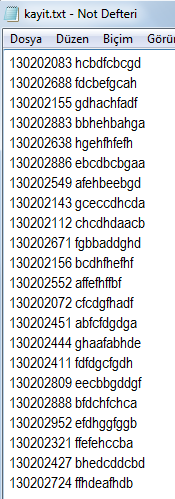
Kayıt dosyasının başarılı bir şekilde oluşturmasının ardından aranacak olan okul numarasının kullanıcı tarafından girilmesi beklenmektedir.

Kullanıcı bir numara girdikten sonra hash fonksiyonlarıyla sırasıyla aramalar yapılmaktadır. Bunlar ekrana basılmaktadır.

Program arama işlemlerine sırayla devam etmektektedir. Arama sonunda da istendiği üzere yapılan işlem sayısı ekranda gösterilmektedir.

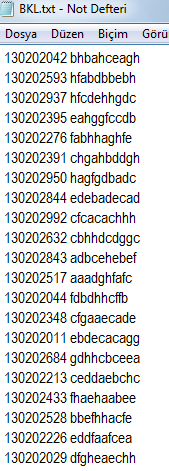
Arama işlemi sonucunda kullanıcı tarafından girilen verinin dosyalar içerisinde bulunamadığı da ekranda gösterilmiştir.

Aşağıdaki görüntülerde ise arama işlemleriyle oluşturulan dosyaların bir kısmı gösterilmiştir.



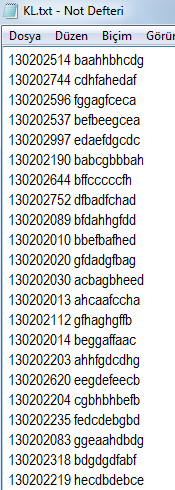
Oluşturulan kayıt dosyası gösterilmektedir.

****

****

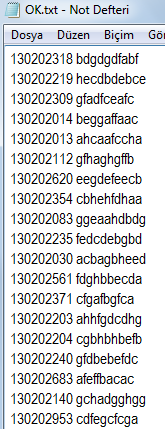
Bölen – kalan yöntemiyle lineer ve ayrık taşmayla oluşturulan dosyalar bu şekildedir.

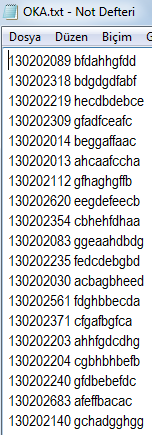




Katlamalı lineer ve katlamalı ayrık taşma yöntemlerinin oluşturduğu dosyalar görselde olduğu gibidir.

Kare ortası ayrık taşma ve kare ortası lineer dosya aşağıda gösterilmiştir.



****

**5.SONUÇLAR**

Proje de kodlama aşamasından daha çok konunun teorik kısmı zorlamıştır. İnternet ortamında yeterli Türkçe kaynak bulunamadığından yabancı kaynaklar üzerinden sonuçları çıkarılmıştır.

Kodlama kısmında katlamalı dosyalar oluşturulurken üçlü gruplama yapılamamıştır. Buna çözüm olarak dokuz basamaklı sayının her bir rakamını kendi başına bir grup olarak düşünüp tekli gruplar şeklinde işlemler yapılmıştır.

**6.KAYNAKÇA**

[1] <https://tr.wikipedia.org/wiki/Dev-C%2B%2B>

[2] <http://erayyelmen.blogspot.com.tr/2014/05/c-hash-fonksiyonlar-ile-dosyaya-arama.html>

[3] <http://erayyelmen.blogspot.com.tr/2014/05/c-hash-fonksiyonlar-ile-dosyaya-arama.html>

[4] <http://ww2.valdosta.edu/~sfares/cs330/cs3410.a.hashing.fa8.html>

[5] Görseller için:  
http://www.csharpnedir.com/articles/read/?id=453