

COSMIC İşlevsel Büyüklük Ölçüm Sonuçlarında Gözlenen Sapmalar Üzerine Bir Deney Çalışması

Erdir Ugan¹

Onur Demirörs²

Barış Özkan³

^{1,2,3}Enformatik Enstitüsü, Orta Doğu Teknik Üniversitesi, Ankara

¹e-posta: erdir@ii.metu.edu.tr

²e-posta: demirors@metu.edu.tr

³e-posta: bozkan@ii.metu.edu.tr

Özetçe

İşlevsel büyüklük ölçümlerinde ölçüm sonuçları, ölçümü gerçekleştiren kişilere bağlı olarak farklılık gösterebilmektedir. Bu çalışmada COSMIC ile gerçekleştirilmiş bir işlevsel büyüklük ölçüm deneyinin sonuçları sunulmaktadır. Çalışmanın amacı, farklı bireylerce yürütülen işlevsel büyüklük ölçümlerinin sonuçları arasındaki farkların analiz edilmesi ve bu sapmalara neden olan unsurların tespit edilmesidir. Çalışma, bir kontrollü deney olarak tasarlanmış ve yürütülmüştür. Makalede sürecin detayları, ölçüm sonuçlarıyla ilgili ham veriler, çalışma sonunda elde edilen bulgular ve analizler sunulmuştur.

1. Giriş

Yazılım projelerinde planlama ancak dayandırıldığı kestirimler kadar başarılı olabilmektedir. Sağlıklı bir boyut kestirimi yapılmadan gerçekleştirilecek her türlü işgücü veya süre kestirimi ise büyük oranda başarısız olacaktır.

Son yıllarda, Allan Albrecht'in fikirlerinden [2] yola çıkan İşlevsel Büyüklük Ölçüm (İBÖ) yöntemleri, yazılım ölçümleme alanında yaygın olarak kabul görmektedir.

İBÖ, yazılım projelerinin erken aşamalarında, gereksinimler belirlendiği zaman ölçümlenebilmesini sağlar. Yöntem, yazılımın işlevselliğini değerlendirdiği için projenin hayata geçiriliş yönteminden bağımsızdır. Bu durum farklı programlama dilleri ve farklı yöntemlerle geliştirilen projelerin incelenmesini sağlayarak kıyas çalışmalarında büyük kolaylık sağlamaktadır.

Yazılım ölçümlemede birçok açıdan avantajlar sağladığı ve yaygın olarak kullanıldığı halde İBÖ'nün de kendine has zorlukları ve sorunları bulunmaktadır. İBÖ'ye yöneltilen eleştirilerin büyük bir kısmı, ölçüm yöntemlerinin en önemli özelliği olan nesnellikten yoksun olması, farklı iki kişinin gerçekleştireceği ölçümlerin farklı sonuçlar vermesi üzerine yoğunlaşmaktadır [17]. Ölçüm sonuçları arasındaki farkların, çeşitli soyutlama seviyelerinde tanımlanabilecek birçok unsurdan kaynaklanması mümkündür.

Bu çalışma, işlevsel büyüklük ölçümü sonuçları arasında farklılıklara ve tutarsızlıklara neden olan unsurları tespit etmek amacıyla gerçekleştirilmiş bir deneyin sonuçlarını sunmaktadır. Deney çalışması, bir lisansüstü ders kapsamında yürütülmüş olup tek bir yazılıma ait gereksinimlerin, farklı kişiler tarafından ölçülmesi yoluyla gerçekleştirilmiştir. Bulgular gruplanarak incelenmiş ve tespit edilen sorunların muhtemel temel nedenleri tartışılmıştır. Çalışmada İBÖ yöntemi olarak COSMIC kullanılmıştır [22].

Çalışmanın ana katkısı, endüstri bünyesinde gerçekleştirilmiş bir vaka çalışması yerine kontrollü bir ortamda yürütülmüş bir deney olmasıdır. Aynı zamanda bu çalışma, bilindiği kadarıyla bu konuda, görece yeni bir İBÖ yöntemi olan COSMIC ile gerçekleştirilmiş ilk çalışmadır.

Bir sonraki bölümde, İBÖ ve COSMIC ile ilgi literatürde yer alan çalışmalar özetlenmektedir. Sonraki bölümde problemin tanımlanması ve deneyin ön çalışmaları anlatılacaktır. Daha sonra, deneyde izlenen yöntem, ve deneyin yürütüldüğü ortam tanımlanacaktır. Çalışmada elde edilen bulgular üç seviyede sunulacaktır. İlk olarak ölçümlerin değerlendirilmesi sonucunda elde edilen hatalar ve tutarsızlıklar hakkında ham istatistiksel veriler verilecektir. İkinci olarak, bu bulgular, hataların doğaları uyarınca gruplanacaktır. Son olarak bu hata gruplarının işaret ettiği temel sorunlar ve bu sorunların kaynakları tartışılacaktır.

2. İlgili Çalışmalar

İşlevselliğe dayalı yazılım ölçümleme fikri ilk olarak 1979'da Albrecht tarafından ortaya atılmıştır [2]. Fonksiyon Nokta Analizi (FNA) adıyla tasarlanan ilk yöntem, daha sonra Gaffney [3] ve Albrecht [4] tarafından geliştirilmiştir. Daha sonraki yıllarda, birçok araştırmacı bu yöntemden yola çıkarak değişik yöntemler geliştirmişlerdir [20].

Yöntemlerin sayısının artması ve kapsamlarının genişlemesi üzerine, ISO/IEC bu ölçüm yöntemlerini standartlaştırmak amacıyla bir çalışma başlatmıştır. Bu çalışma sonucunda "14143: Information Technology - Software Measurement" adlı İBÖ standartları yayınlanmıştır [6] - [11]. Şu an itibarıyla, Mark II FPA[14], IFPUG FPA [13], COSMIC FFP [12], NESMA FSM [15], ve FISMA [16] olmak üzere beş yöntem ISO tarafından onaylanmıştır.

İBÖ yöntemlerinin güvenilirliği en baştan beri birçok araştırmaya konu olmuştur. Low ve Jeffrey, [19] büyüklük ölçümü ve yazılım geliştirme alanlarında farklı deneyim seviyelerine sahip katılımcılar ile gerçekleştirdikleri deneysel çalışmada Fonksiyon Nokta ölçüm sonuçlarının tekrar üretilebilirliğini araştırmıştır. Çalışmanın sonuçları, tecrübeli katılımcıların gerçek ölçüm sonuçlarına daha yakın sonuçlar elde ettiklerini göstermiştir. Bir erken ölçüm tekniği olarak, FN'nin Kod Satır Sayısı ölçüsü ile karşılaştırıldığında daha doğru kestirim sonuçları verdiği vurgulanmıştır.

1992 yılında Kemerer ve Porter [17] FN yönteminin güvenilirliğini etkileyen temel nedenleri belirleyebilmek amacıyla yaptıkları çalışma sonucunda hata tiplerinin [17] küçük bir oranının büyük etkilere neden olduğu ortaya koymuşlardır. Ölçüm sonuçlarının güvenilirliğini artırmak için

açıklanan hata tiplerini engelleyecek çözümler ortaya koyulması, IFPUG FPA yönteminin kurallarını daha net anlaşılacak biçimde yayınlaması ve değişkenliğin nedenlerini sürekli araştırarak bir program oluşturulması önerilerini getirmişlerdir.

Kemerer, E-R yaklaşımı ile işlevsel büyüklük ölçme ile standart IFPUG FPA yöntemi ile işlevsel büyüklük ölçme tekniklerini karşılaştırmak amacıyla 1993 yılında bir çalışma yapmıştır [18]. Çalışmanın istatistiksel sonuçlarına göre; FP ölçümü yapan kişiler benzer sonuçlara ulaşmaktadırlar. E-R yöntemi ile standart IFPUG karşılaştırıldığında ise iki yöntemin birbirine üstünlüğü gözlenmemiştir.

2004 yılında, Silvia, İBÖ yöntemlerinin nesne yönelimli yazılım sistemleri için tekrar edilebilirliği ve doğruluğunu araştıran bir çalışma yapmışlardır [1]. Çalışmada IFPUG FPA ve Object Oriented Method Function Point (OOmFP) yöntemleri değerlendirilmiştir. 22 katılımcı ile gerçekleştirilen kontrollü deneyin sonuçları OOmFP yönteminin IFPUG FPA'ya kıyasla daha tutarlı ve doğru olduğunu göstermiştir.

2008 yılında Oktay Türetken'in yürüttüğü bir çalışmada varlık genelleme kavramının COSMIC ve IFPUG işlevsel büyüklük ölçme yöntemleri ile bir arada kullanılmasıyla orta çıkan farklı ölçüm sonuçlarının etkileri tartışılmıştır [23]. Türetken'in bir diğer çalışması [24], COSMIC, IFPUG ve MkII ile ölçüm sırasında karşılaşılan zorluklara işaret etmektedir. Bu çalışmada farklı kişilerin kuralları bozmayacak varsayımlar ve algılamalar nedeniyle aynı proje için farklı ölçüm sonuçlarına erişebilecekleri görülmüştür.

İBÖ yöntemlerinin güvenilirliği üzerine yürütülen çalışmaların neredeyse tümü girdi olarak endüstriyel verileri ve vaka analizlerini kullanmıştır. Endüstriyel veriler veya karşılaştırma çalışmaları için hazırlanan veri kümelerinin kullanımı yüksek ölçüde veri kirliliğine yol açmaktadır. Bu verilerin toplandığı projelerin tipleri, boyutları ve sektörleri arasındaki farklar; kurumların boyutları, yetkinikleri arasındaki farklar ve ölçümleri gerçekleştiren kişilerin deneyimlerindeki değişimler gibi etmenler; ölçümlerde yapılan ortak hataların ve bu hataların nedenlerinin belirlenmesini çok zorlaştırmaktadır. Ayrıca, bahsedilen çalışmaların sonuçları, uygulanan yöntemlerin ayrıntıları bilinemediğinden yeniden üretilememekte ve doğrulanmamaktadır.

3. Çalışmanın Amacı

Çalışmanın ana amacı, işlevsel büyüklük kestirimlerinde hataların sıklıkla gerçekleştiği noktaları tespit etmek ve farklı kişilerce gerçekleştirilen ölçümler arasındaki tutarsızlıkların nedenlerini araştırmaktır.

Çalışma öncesinde, sıklıkla karşılaşılan hatalar ve bu hataların muhtemel nedenleri hakkında kesin yargılar bulunmamasına rağmen, ölçümler arasında sapmaya neden olabilecek unsurlar ve bunların sınıflandırılması hakkında bazı öngörülerden bahsedilebilir.

Hatalı ölçüm; ölçümü gerçekleştiren kişinin, kullanılan İBÖ yönteminin kurallarını gerektiği şekilde uygulayamaması durumudur. Bu durum eğitim eksikliğinden veya yöntemdeki

yer alan bazı muğlak veya anlaşılması zor kurallardan kaynaklanabilir.

Ölçümü gerçekleştiren kişi tarafından benimsenen varsayım ve kabuller; gereksinimlerde belirtilmeyen bazı işlevselliklerin kişisel deneyime veya örnek uygulamalara dayanarak var kabul edilerek ölçüm sırasında eklenmesi durumudur.

Yorum farklılıkları; bir problemin farklı kişilerce farklı analiz edilmesi durumudur. Farklı analizlere dayanan ölçümler, ölçüm yöntemi tamamen doğru olarak uygulanırsa bile, farklı sonuçlar verebilir.

Çalışmada, farklı kişilerin gerçekleştirdiği ölçümler arasındaki sapma miktarları gösterilmeye çalışılmıştır. Bu sapma, "eksik fonksiyonel süreçler", "fazla fonksiyonel süreçler" ve "veri hareketlerindeki sapmalar" gibi değişik boyutlarıyla incelenmiştir.

Ayrıca, çalışma sırasında oluşturulan veri kümesi, kontrollü bir ortamda oluşturulmasından dolayı, ileride gerçekleştirilebilecek benzeri çalışmalar için bir temel oluşturabileceğinden, çalışmanın bir yan ürünü olarak kabul edilebilir.

4. Deneyin Tasarımı

Çalışmada yürütülen deney, bir kontrollü deneyin gereklerini karşılayacak şekilde tasarlanmıştır. Bu nedenle, bazı ön çalışmalar planlanmıştır.

İlk olarak, katılımcıların İBÖ ve COSMIC hakkında yeterli bilgi seviyesine sahip olmaları gerektiğinden, bu konularda eğitim verilmesi planlanmıştır. Eğitim, önce yazılım ölçümleme ve işlevsel büyüklük ölçüm kavramları hakkında sınıf eğitimleri şeklinde olacak, ardından COSMIC yöntemi hakkında özel eğitimler verilecektir.

Ardından, katılımcıların konu hakkında deneyim kazanması amacıyla bir pilot uygulama gerçekleştirilecektir. Katılımcılar örnek bir yazılım projesini ölçümleyecek, ölçüm sonuçları konunun uzmanlarınca değerlendirilecek ve katılımcılara birebir geribildirim verilecektir.

Katılımcılar hakkında bilgi toplamak ve ölçümü gerçekleştiren kişilerin ölçümler üzerindeki etkilerini araştırabilmek amacıyla, katılımcıların yazılım geliştirme ve yazılım ölçümleme konularındaki bilgi seviyelerini ölçecek bir anket hazırlanacak ve uygulanacaktır.

Deney için, basit bir problem oluşturulacak ve bu problem için önerilen bir yazılım çözümü betimlenecektir. Katılımcılara birkaç madde gereksinim verilecek ve ölçümleme bu gereksinimleri üzerinden gerçekleştirilecektir. Problem basit olmakla beraber, gerçek hayatta karşılaşılabilecek türden bir problem olacaktır.

Ölçümleme sırasında karşılaşılan sorunları etkin bir biçimde belirleyebilmek amacıyla, problem tanımı, basit, bileşik, zayıf ve güçlü varlık tiplerinden örnekler ve bunların ilişkilerini içermelidir.

Önerilecek yazılım çözümünün gereksinimleri, basit sorgular ve işlemler; bileşik varlıklar ve bunların alt varlıkları üzerine işlemler; örtülü işlemler ve basamaklı işlemler içermelidir.

Grup içinde tartışmaların ve fikir alışverişlerinin önlenmesi amacıyla deneyin üç saat içerisinde tamamlanması planlanmıştır.

Ölçüm sonuçları, deneyde verilen gereksinimler için doğru fonksiyonel süreç, ilgi nesnesi ve veri hareket tanımlamalarını içeren bir anahtar kullanılarak değerlendirilecektir. Bu anahtar, COSMIC İBÖ üzerinde akademik çalışma yürüten ve endüstriyel deneyime sahip uzmanlar tarafından ortaklaşa oluşturulacaktır.

Sonuçlar yinelemeli bir şekilde değerlendirilecektir. İlk olarak, benzer hatalar tanımlanacaktır. Sonra, bu hatalar doğalarına göre gruplanacak, son olarak da, bu hata gruplarına yol açan muhtemel nedenler tartışılacaktır. Deneyin tasarımıyla ilgili; anahtar çözüm, veri toplamada kullanılan formlar, pilot uygulamada ve deneyde kullanılan problem tanımları gibi detaylar ilgili teknik raporda bulunabilir [25].

5. Deneyin Yürütülüşü

5.1. Hazırlık

Çalışma Ortadoğu Teknik Üniversitesi Enformatik Enstitüsü bünyesindeki "Yazılım Proje Yönetimi" adlı lisansüstü dersin bir parçası olarak yürütülmüştür. Öğrenciler yazılım ölçümleme, İBÖ ve COSMIC üzerine toplam 9 ders saati eğitim almışlardır.

Daha sonra, pilot uygulama gerçekleştirilmiştir. Katılımcılara deneyde kullanılacak probleme benzer bir problem verilmiş ve ödev olarak bu gereksinimleri COSMIC ile ölçmeleri istenmiştir. Sonuçlar toplanmış ve değerlendirilmiş, doğru ölçüm şekli sınıfta tartışılmış, katılımcılara ölçümlerindeki hatalarla ilgili birebir geribildirim verilmiştir. Pilot uygulamanın tartışıldığı oturum 3 saat sürmüştür. Eğitim ve örnek uygulama sonunda katılımcıların COSMIC hakkında belirli bir bilgi seviyesine ulaşması sağlanmıştır.

5.2. Örnek Problem

Çalışmada verilen problem, filmler hakkında bilgiler içeren bir veritabanı üzerinde çalışacak bir bilgi sistemini tanımlamaktadır. Betimlenen yazılım çözümü, veri girişi, veri güncelleme ve çeşitli sorgular içermektedir.

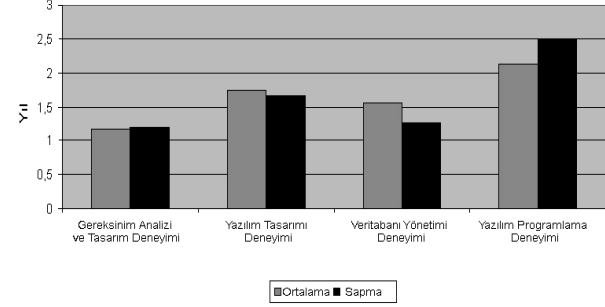
Katılımcılara ölçmeleri için her biri alt gereksinimler içeren 6 gereksinim maddesi verilmiştir. Bu gereksinimler, sorgulama, listeleme, güncelleme, silme, basamaklı silme, kısmî güncelleme ve kısmî silme içeren işlemler talep etmektedir.

Problem tanımı ve gereksinimlerin yanısıra, gerçekleştirilecek kısa bir analiz sonucu elde edilebilecek basit bir varlık ilişki diyagramı da katılımcılara verilmiştir. Problem, basit, bileşik, zayıf ve güçlü varlıklar içermektedir.

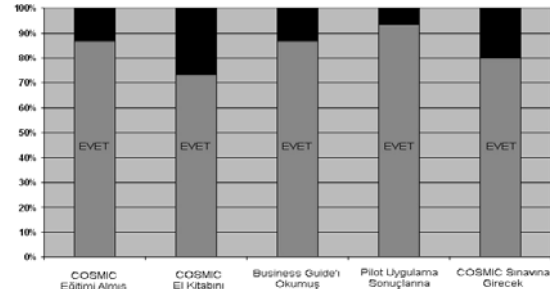
Tanımlanan gereksinimlere göre oluşturulacak yazılım 12 fonksiyonel süreçten oluşacak olup, boyutu 82 CFP (Cosmic Functional Points) olarak ölçülmektedir.

5.3. Katılımcılar

Deney sınıf ortamında 15 katılımcının toplu katılımıyla gerçekleştirilmiştir. Katılımcılara önce deney ve uygulanacak yöntem hakkında önbilgiler verilmiş, daha sonra yazılım ölçümleme konusundaki deneyimleri, iş deneyimleri ve COSMIC hakkındaki bilgi seviyeleri hakkında bilgi toplamak amacıyla hazırlanmış anketler dağıtılmış ve doldurmaları istenmiştir. Katılımcıların ilgili yazılım geliştirme alanlarında ortalama deneyim seviyeleri şekil 1'de yıl cinsinden gösterilmektedir.



Şekil 1: Ortalama katılımcı deneyim seviyesi.



Şekil 2: Katılımcıların COSMIC bilgi seviyeleri.

Katılımcıların ölçümleri ellerinden gelen en iyi biçimde gerçekleştirmelerini sağlamak amacıyla, ölçümlerdeki başarı oranları, ilgili dersten alacakları başarı notuyla ilişkilendirilmiştir.

Ölçümden önce ve sonraki tüm tartışmalar videoya kaydedilmiştir. Bu görüntüler, daha sonra, ölçümlerdeki hataların nedenlerini belirlemek amacıyla incelenmiştir.

5.4. Ölçüm Sonuçları

Fonksiyonel süreçlerin, ilgi nesnelere, ve veri hareketlerinin ne şekilde belirlenmesi gerektiğini belirten bir çözüm anahtarı hazırlanmış ve katılımcıların ölçüm sonuçları bu anahtara göre değerlendirilmiştir. Hem tespit edilememiş hem de fazladan eklenmiş fonksiyonel süreçler ve veri hareketleri ölçüm hataları olarak kabul edilmiştir. Her bir katılımcının doğru ölçümden ne kadar saptığı belirlenmiş, sapmalar, eksik ve fazla Fonksiyonel Süreç (FS) ve Veri Hareketi (VH) olarak verilmiştir. Katılımcıların ölçümlerindeki sapmalar ayrı ayrı ve ortalama olarak tablo 1'de görülmektedir.

Tablo 1. Ölçüm Sonuçları Özeti

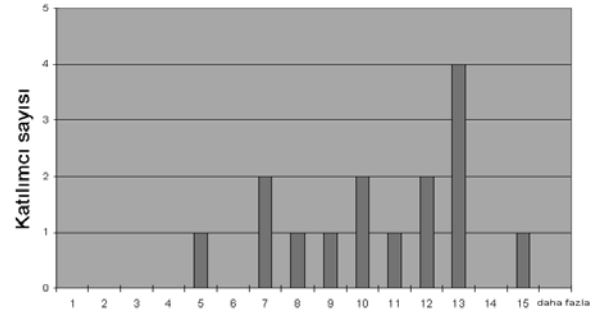
Katılımcı No.	Toplam FS	Doğru FS	% Doğru FS	Fazla FS	Toplam VH	Doğru VH	% Doğru VH	Fazla VH
1	5	5	41,67%	0	49	35	42,68%	14
2	12	11	91,67%	1	82	56	68,29%	26
3	13	11	91,67%	2	78	66	80,49%	12
4	7	7	58,33%	0	52	42	51,22%	10
5	7	7	58,33%	0	52	46	56,10%	6
6	10	10	83,33%	0	61	57	69,51%	4
7	13	12	100,00%	1	53	48	58,54%	5
8	10	9	75,00%	1	35	15	18,29%	20
9	11	10	83,33%	1	49	47	57,32%	2
10	8	8	66,67%	0	46	37	45,12%	9
11	13	10	83,33%	3	67	44	53,66%	23
12	12	10	83,33%	2	83	52	63,41%	31
13	13	12	100,00%	1	75	57	69,51%	18
14	15	12	100,00%	3	93	65	79,27%	28
15	9	6	50,00%	3	28	21	25,61%	7
Ortalama	10,53	9,33	77,78%	1,20	60,20	45,87	55,93%	14,33

Ölçüm sonuçları çözüm anahtarına göre değerlendirilirken bazı kurallar izlenmiştir.

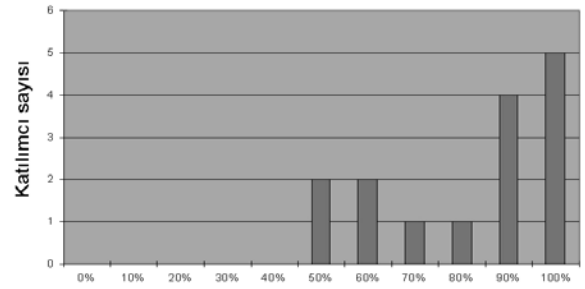
- Doğru tespit edilmiş bir FS içindeki fazla VH'ler, "fazla VH" olarak değerlendirilmiştir.
- Fazla FS'lerin içindeki VH'ler "fazla VH" olarak değerlendirilmiştir.
- İki ayrı FS'nin (ve onlara ait VH'lerin) birleştirilip tek bir FS gibi sayıldığı durumlar VH hatası olarak kabul edilmemiş, sadece eksik FS dikkate alınmıştır.
- Bir bileşik ilgi nesnesi ile ilgili veri hareketleri tek bir VH olarak sayıldığında, işlemden sayılması gereken VH'lerin sadece biri tespit edilmiş, diğerleri tespit edilememiş kabul edilmiştir.

Ölçümlerin güvenilirliği veya doğruluğu için tek bir kriter belirlenmemiştir. Bunun yerine, katılımcıların belirlediği Fonksiyonel Süreçler ve Veri Hareketlerinde, toplam belirlenen, doğru olarak belirlenen ve fazladan eklenenler tespit edilmiştir. Sonuçlar, ölçümler arasındaki farkları vurgulamak amacıyla histogram grafikler halinde sunulmuştur.

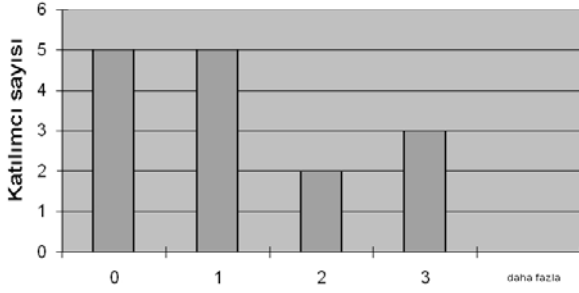
Benzeri bir analizi ilgi nesnelerinin tespitiyle ilgili olarak yapmak mümkün olmamış, ilgi nesnelerinin belirlenmesi ile ilgili sorunlar sonraki bölümde tartışılmıştır.



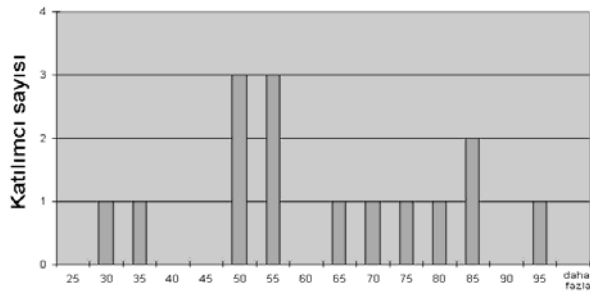
Şekil 3: Toplam tanımlanan FS sayısı.



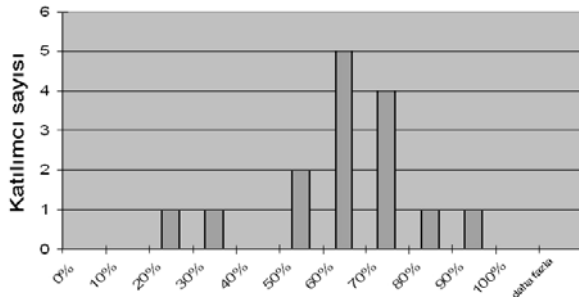
Şekil 4: Doğru tanımlanan FS oranı.



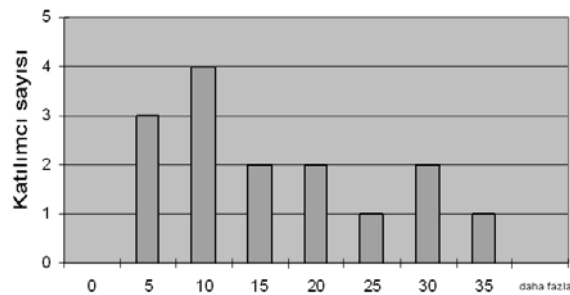
Şekil 5: Fazladan tanımlanmış FS Sayısı.



Şekil 6: Toplam tanımlanan VH Sayısı.



Şekil 7: Doğru tanımlanan VH oranı.



Şekil 8: Fazladan tanımlanmış VH sayısı.

6. Çalışmanın Kısıtları

Yürütülen çalışmada bazı kısıtlar bulunmaktadır.

Katılımcı sayısı istatistiksel olarak anlamlı sonuçlar vermek için yeterli değildir. Çalışmanın bu anlamdaki eksikliğini gidermek amacıyla, aynı deneyin bir sonraki akademik dönemde başka katılımcılarla tekrarlanması planlanmıştır. Bu çalışma sonunda elde edilen sonuçlar yayınlanacak ve ilgili teknik rapor [25] güncellenecektir.

Katılımcılar İBÖ ve COSMIC üzerine 9 saat eğitim almışlardır. Bu eğitimin yeterliliğinin nicel olarak değerlendirilmesi mümkün olmamıştır. Ancak, deneyimler, verilen eğitimin, sektördeki şirketlerin projelerini ölçümlemekle görevlendirdiği personele sağladığı eğitimlerden eksik olmadığını göstermektedir. Ayrıca, katılımcıların 2'si haricinde hepsi deneyden 5 gün sonra COSMIC sertifikasyon sınavına girerek COSMICON (The Common Software Measurement International Consortium) sertifikası almaya hak kazanmışlardır.

Ölçümlenen gereksinimler sayıca azdır ve yapay olarak geliştirilmiştir. Bulguların daha büyük projeler için de geçerli olup olmayacağı bilinmemektedir. Ancak, daha fazla sayıda gereksinim kullanıldığı takdirde, deney için amaçlanan izolasyonun sağlanması mümkün olmayacağından gereksinim sayısı sınırlı tutulmuştur. Proje boyutu büyüdükçe büyüklük ölçüm süreci daha geniş bir zaman dilimine yayılacaktır. Bu durumun, tek bir kişinin bile değişik zamanlardaki ölçümleri arasında tutarsızlığa yol açması beklenir. Bu durum da çalışmada hedeflenen kontrollü deney ortamının bozulmasına ve ölçüm sorunlarının araştırılmasında zorluklara neden olacaktır.

Bazı ölçümlerin çözüm anahtarları ile çeliştikleri ve verilen gereksinimleri tam olarak ölçemedikleri halde kendi içlerinde tutarlı oldukları gözlemlenmiştir. Bu ölçümler, problemde tanımlanan varlıkların ve bunların ilişkilerinin farklı yorumlanması sonucunda elde edilmiştir. Bu ölçümler, problem tanımı ve gereksinimlerle beraber bir varlık ilişki diyagramı da verildiğinden ve [21]'de yer alan ilkelerin uygulanması gerektiğinden hatalı kabul edilmiştir.

7. Bulgular

Ölçümlerde yapılan hatalar ve sonuçlar arasındaki tutarsızlıklar, türlerine göre; ilgi nesnelere tanımlanmasıyla ilgili olanlar, fonksiyonel süreçlerin tanımlanmasıyla ilgili olanlar, varlıkların ve ilişkilerinin farklı yorumlanmasından kaynaklananlar, ölçücülerin yapmış olduğu farklı varsayımlardan kaynaklananlar ve COSMIC yönteminin kurallarının eksik anlaşılmasından kaynaklananlar olmak üzere gruplara ayrılmıştır. Deneyde kullanılan probleme özgü kısıtların gözden kaçırılması, ölçülmemiş gereksinimler ve satır atlama gibi basit hatalar değerlendirmeye alınmamıştır.

7.1. İlgi Nesnelere Belirlenmesi

Parametre tablolarında yer alan veriler bir ilgi nesnesi olarak değerlendirilmiş ve bu nedenle bu veriler üzerinde gerçekleştirilen işlemler için veri hareketleri tanımlanmıştır. Katılımcıların %70'inin ölçümlerinde bu durumdan kaynaklanan fazla veri hareketleri gözlemlenmiştir.

En sık rastlanan hataların ikisi ise ilgi nesneleri arasındaki hiyerarşi ile ilgilidir.

- Katılımcıların hiçbirinin, ilgi nesnelere, sorgu ve benzeri yöntemlerle oluşturulan alt grupları üzerinde gerçekleştirilen işlemleri doğru değerlendiremediği gözlemlenmiştir. Alt gruplar yerine ana ilgi nesnelere değerlendirilmeye alınmıştır. Örnek: "Oyunculuk yapan kişiler" sorgusunda ilgi nesnesi olarak "kişi" nesnesinin kabul edilmesi.
- Katılımcılar, bileşik ilgi nesnesi içeren ölçümlerde, ya sadece ana ilgi nesnesini ya da sadece bu nesnenin alt nesnelere kullanmışlardır. Şöyle ki; bir nesneyi ana nesne olarak algılayan bir kişi, bu nesnenin alt grupları üzerinde işlem içerenler de dahil olmak üzere bütün fonksiyonel süreçlerde bu ana nesneyi kullanmıştır. Bu durum veri hareketi sayısının eksik tespit edilmesine yol açmıştır. Bu davranış katılımcıların %50'sinde gözlemlenmiştir.
- Benzeri şekilde, bir nesnenin alt parçalarını tespit eden bir kişi sadece ana nesne üzerinde gerçekleştirilen işlemlerde dahi tüm alt nesnelere için veri hareketleri tanımlamıştır. Bu durum da veri hareketi sayısının fazla tespit edilmesine neden olmaktadır. Bu davranış katılımcıların %30'unda gözlemlenmiştir.

7.2. FS'lerin Belirlenmesi

Fonksiyonel süreçler belirlenirken, birden çok FS birleştirilerek tek bir FS gibi sayılmış veya bazı FS'ler hatalı bir şekilde birden çok parçaya bölünmüşlerdir.

Katılımcıların %70'i "sorgulama" ve "detay listeme" işlemlerini tek bir FS'de birleştirmişlerdir. Katılımcıların, gereksinim metninde ard arda belirtilmiş FS'leri birleştirme eğiliminde oldukları açık bir şekilde gözlemlenmiştir.

Koşul içeren işlemlerde, bir FS'nin her biri ayrı bir koşulu tanımlayacak şekilde ayrı FS'lere bölündüğü gözlemlenmiştir.

FS'leri tetikleyen olaylar neredeyse tüm katılımcılar tarafından farklı tanımlanmıştır. Ancak yöntemde tetikleyici olaylar tek bir veri hareketi olarak değerlendirildiğinden, ölçümlerin sayısal sonuçlarını etkilememiştir.

7.3. Yorum Farklılıkları

Hatalı olarak değerlendirilen bazı ölçümlerin kendi içlerinde aslen tutarlı olabilecekleri gözlemlenmiştir. Bu ölçümler, problemdeki varlıklar ve bunların ilişkileri farklı yorumlanarak gerçekleştirilmiştir. Bu ölçümler kendi içlerinde tutarlı olmalarına rağmen, katılımcıların kullandığı ortak bir varlık ilişkisi diyagramının bulunması ve ölçüm sırasında hatalı bir şekilde bu diyagramdan sapılması, bu sonuçların hatalı kabul edilmesine neden olmuştur.

7.4. Varsayımlar

İşlevlerin yeniden kullanımı hakkındaki varsayımlar : Belli bir seviyede programlama bakış açısına sahip katılımcıların, başka bir FS'de tanımladıkları işlevleri diğer FS'lerde yeniden kullanma eğiliminde oldukları gözlemlenmiştir. Bu duruma en

çok, "silme/güncelleme öncesi listeleme" ve "silme/güncelleme öncesi getirme" işlemleri içinde yer alan "listeme" ve "getirme" işlevlerinin yeniden kullanımı şeklinde rastlanmıştır.

Hata ve onay mesajlarıyla ilgili varsayımlar: Katılımcıların %85'i gereksinim metninde açıkça belirtilmedikçe, hata veya onay mesajlarından doğacak "Çıkış" VH'leri tanımlamamışlardır. Ancak kalan %15'in bu tür mesajların beklenebileceği her durumda varlıklarını kabul ederek ilgili veri hareketlerini tanımladıkları gözlemlenmiştir.

Katılımcıların %40'ı, "silme" işlemi öncesinde ek bir "silme işlemi öncesi getirme" fonksiyonel sürecinin bulunduğunu varsayımlamışlardır. Bu varsayımda bulunanların yarısı ise, bu amaç için "listeme" fonksiyonel sürecini yeniden kullanmışlardır.

7.5. Ölçüm Kuralları

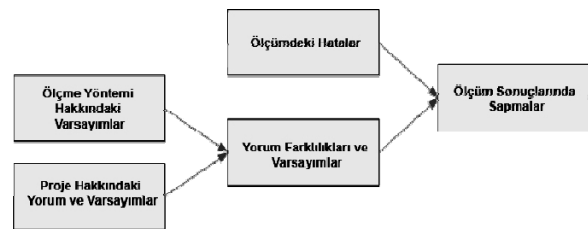
Açılan listelerin doldurulması sırasında, katılımcıların %72'si "Okuma", %86'sı ise "Çıkış" işlemlerini dikkate almamıştır.

Katmanların algılanmasındaki tutarsızlıklar: Bazı katılımcılar güncelleme işlemleri dahilindeki yer alan "Yazma" hareketleri öncesinde bir "Okuma" hareketinin bulunduğunu varsayımlamışlardır. Veritabanı tasarlama ve yönetimi konularında deneyim sahibi katılımcılar, veritabanı yönetimi dahilinde gerçekleştirilen işlemleri de dikkate alarak bu katmandaki veri hareketlerini de sayma eğilimi göstermektedir. Ölçümlenen katman uygulama katmanı olup, diğer katmanlardaki veya katmanlar arasındaki veri hareketleri ölçümlere dahil edilmemelidir.

Basamaklı silme işlemleri ile ilgili sorunlar: Katılımcıların %80'i bir ilgi nesnesi silinirken onunla ilişkili nesnelere de getirilmesini sağlayacak "Okuma" hareketlerini dikkate almamıştır. Buna rağmen, katılımcıların yine %80'i her ilişkili nesnenin silinmesi için gereken "Yazma" hareketlerini tanımlayabilmiştir. Katılımcıların %10'u tüm ilişkili nesnelere için tek bir "Okuma" hareketi tanımlamış, %10'u ise sorgulama işlevine sahip FS'lerin tekrar kullanıldığını varsayımlamıştır.

8. Sonuçlar

COSMIC ölçümlerinde gözlenen sapmanın iki temel boyutu bulunmaktadır. İlki, ölçüm yönteminin kurallarının uygulanmasındaki hatalar, ikincisi ise, bu kurallar doğru uygulansa bile kişilerin kendi yorum ve varsayımlarından kaynaklanan farklılıklardır. Ölçüm sapmalarına yol açan unsurlar Şekil 9'da betimlenmiştir.



Şekil 8: Ölçüm sonuçlarında sapmaya yol açan unsurlar.

Ölçüm yönteminin kurallarının yanlış uygulanması, yetersiz eğitim veya deneyim eksikliğine bağlanabilir.

Yöntem için standartlaştırılmış eğitim materyallerinin geliştirilmesinin özellikle deneyimsiz kişilerce yürütülecek ölçümlerde hata oranının azalmasına yardımcı olacağı söylenebilir. Bu minvalde, sıkça rastlanan işlemler hakkında daha çok örnek ve açıklama içeren uygulama kılavuzlarının yayınlanması yararlı olacaktır.

Farklı yorum ve varsayımlara yol açan unsurlar da kendi aralarında; ölçüm yöntemi hakkındaki ve ölçülen proje hakkındaki yorum ve varsayım farkları olarak iki kategoride incelenebilir.

Önceki başlıkta bahsedilen bulgulardan; bazı katılımcıların varsaydığı hata ve onay mesajlarına bağlı “Çıkış” hareketleri veya “güncelleme/silme” işlemlerinden önce bir “getirme” işleminin bulunduğu varsayılması gibi tutarsızlıklar, ölçüm yönteminin uygulanışı ile ilgili varsayımlara birer örnektir.

Yöntem elkitabında, bu gibi durumlar için açık ve net açıklamalar bulunduğu takdirde bu tür tutarsızlıklarda önemli bir düşüş olacağı söylenebilir. [21]'de bu yönde örnekler bulunmaktadır. Ancak, olabilecek her durum için, bu kaynaktaki şekli ile bir örnek sunulması mümkün değildir. Bu nedenle, bu tip durumlarda genel olarak izlenecek yaklaşım, yöntemin ana elkitabında tanımlanmalıdır. Aksi takdirde açıkça örneklendirilmeyen her durum ölçümlerde tutarsızlıklara neden olacaktır. Tanımlanacak yaklaşımlar, genel geçer olmakla beraber, yoruma yer bırakmayacak şekilde verilmelidir.

Yine bir önceki başlıkta bahsedilen, katmanların algılanmasındaki tutarsızlıklar, işlevlerin yeniden kullanımı hakkındaki varsayımlar ve varlıklar ile bunların ilişkilerinin tanımlanmasındaki farklılıklar, ölçülen proje hakkındaki yorum ve varsayım farklarına örnektir. Ölçümü gerçekleştiren kişilerin, ölçülen yazılım katman veya katmanları, hangi işlevlerin yeniden kullanılacağı ve problemin içerdiği varlıklar arasındaki ilişkiler üzerinde ortak bir anlayışa sahip olmaları, bu tip durumlardan kaynaklanan sapmaları büyük ölçüde azaltacaktır. Yöntemin direk olarak tek bir basamakta uygulanması yerine, bir takım aşamalardan oluşan bir ölçüm sürecinin tanımlanması ile bu sağlanabilir. Bu süreç, ölçüm kapsamının belirlenmesini sağlayacak yöntemler ve problemin ölçüme yönelik toplu bir analizinin yapılmasını sağlayacak etkinlikler içermelidir. Çalışmada da görüldüğü üzere, varlık ilişki diyagramı gibi alışılmış yazılım analiz yöntemleri, bir ekipçe gerçekleştirilecek ölçümlerde tutarlılığın sağlanması için yetersiz kalabilmektedir. Özel olarak yazılım ölçümüne yönelik analiz yöntemleri içeren bir ölçüm süreci bu tip sapmalarda azalma sağlayacaktır.

Bu çalışmada, tek bir yazılım problemi, kontrollü bir ortamda, farklı kişilerce, COSMIC İBÖ yöntemi kullanılarak ölçülmüştür. Ölçüm sonuçlarında farklılıklara neden olan unsurlar, farklı projelerin ölçülmesinden kaynaklanabilecek veri kirliliğinden etkilenmeden tespit edilmiştir. Çalışma sonucunda, yöntemin kendisinde ve uygulanmasında gerçekleştirilecek bazı iyileştirmelerle bu unsurlardan kaynaklanan sapmaların önüne geçmenin mümkün olduğu görülmüştür.

9. Teşekkür

Deneyde kullanılan örnek problem tanımı ve gereksinimlerin hazırlanmasındaki katkılarından dolayı Dr. Oktay Türetken'e, ilgili çalışmalar ve literatür taramasındaki katkılarından dolayı Özden Özcan Top'a teşekkür ederiz.

10. Kaynakça

- [1] Abrahao S, Poels G, Pastor O. *Assessing the reproducibility and accuracy of functional size measurement methods through experimentation*. ISESE 04. 2004
- [2] Albrecht, A. *Measuring Application Development Productivity*. Joint SHARE / GUIDE / IBM Application Development Symposium, 83, 92. 1979
- [3] Albrecht, A.J. and Gaffney J.E. *Software Function, Source Lines of Code, and Development Effort Prediction: A Software Science Validation*, IEEE Transactions on Software Engineering, Cilt 9, Sayı 6, Sayfa 639-648, Kasım. 1983.
- [4] Albrecht, A.J., *AD/M Productivity Measurement and Estimate Validation*, IBM Corporate Information Systems, IBM Corp., Purchase, N.Y., Mayıs 1984.
- [5] <http://www.gelog.etsmtl.ca/cosmic-ffp/cert.html>
- [6] ISO/IEC 14143-1: Information Technology – Software Measurement - Functional Size Measurement - Part 1: Definition of Concepts. 1998. 2007'de güncellenmiştir.
- [7] ISO/IEC 14143-2: Information Technology – Software Measurement - Functional Size Measurement - Part 2: Conformity Evaluation of Software Size Measurement Methods, ISO/IEC 14143-1:1998, 2002.
- [8] ISO/IEC TR 14143-3: Information Technology – Software Measurement - Functional Size Measurement - Part 3: Verification of Functional Size Measurement Methods, 2003.
- [9] ISO/IEC TR 14143-4: Information Technology – Software Measurement - Functional Size Measurement - Part 4: Reference Model, 2002.
- [10] ISO/IEC TR 14143-5: Information Technology – Software Measurement - Functional Size Measurement - Part 5: Determination of Functional Domains for Use with Functional Size Measurement, 2004.
- [11] ISO/IEC 14143-6: Guide for the Use of ISO/IEC 14143 and related International Standards, 2006.
- [12] ISO/IEC 19761: COSMIC Full Function Points Measurement Manual, v. 2.2, 2003.
- [13] ISO/IEC 20926: Software Engineering - IFPUG 4.1 Unadjusted FSM Method - Counting Practices Manual, 2003.
- [14] ISO/IEC 20968: Software Engineering - MkII Function Point Analysis - Counting Practices Manual, 2002.
- [15] ISO/IEC 24570: Software Engineering - NESMA Functional Size Measurement Method v.2.1 - Definitions and counting guidelines for the application of Function Point Analysis, 2005.
- [16] ISO/IEC 29881: Information Technology - Software and systems engineering. FISMA 1.1 functional size measurement method, 2008.
- [17] Kemerer, C. F. *Reliability of function points measurement: a field experiment* Comm. ACM Cilt 36 Sayı 2. 1993.

4. ULUSAL YAZILIM MÜHENDİSLİĞİ SEMPOZYUMU - UYMS'09

- [18] Kemerer, C.F., Porter, B.S. *Improving the Reliability of Function Point Measurement: An Empirical Study*, IEEE Transactions on Software Engineering, Cilt. 18, Sayı. 11, Sayfa. 1011-1024, Kasım 1992.
- [19] Low G.C. and Jeffery D.R. *Function points in the estimation and evaluation of the software process*. IEEE Transactions on Software Engineering, Cilt 16, Sayı 1, Sayfa : 64 – 71, Ocak 1990.
- [20] Symons, C. *Come Back Function Point Analysis (Modernized) – All is Forgiven!*, 4th European Conf. on Software Measurement and ICT Control, FESMA-DASMA, 2001.
- [21] The Common Software Measurement International Consortium (COSMIC). *Guideline for Sizing Business Applications Software Using COSMIC-FFP, Version 1.0*, 2005.
- [22] The Common Software Measurement International Consortium (COSMIC). *COSMIC Method, Version 3.0, Measurement Manual*, 2007.
- [23] Türetken, O., Demirörs, O., Özcan Top, O., & Özkan, B., *The Effect of Entity Generalization on Software Functional Sizing: A Case Study*. A. Jedlitschka and O. Salo (Eds.): PROFES 2008, LNCS 5089, Sayfa. 105–116,. Springer-Verlag Berlin Heidelberg, 2008.
- [24] Türetken, O., Özcan Top, Ö., Özkan, B., Demirörs, O. *The Impact of Individual Assumptions on Functional Size Measurement*. R. Dumke et al. (Eds.): IWSM / MetriKon / Mensura 2008, LNCS 5338, pp. 164–178. Springer Berlin Heidelberg, 2008.
- [25] Urgan, E., Demirörs, O., Özkan, B. ve Özcan Top, Ö. *An Experiment on COSMIC FFP Measurement Variances*. Teknik Rapor. Enformatik Enstitüsü. Orta Doğu Teknik Üniversitesi, 2009.