



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

Centro Nacional de Investigación
y Desarrollo Tecnológico

Tesis de Maestría

**Evaluación Experimental para Determinar la Relación
entre las Métricas de Cohesión y Acoplamiento en
Arquitecturas Orientadas a Objetos**

Presentada por
Ing. Heidi Barrera Monje

como requisito para la obtención del grado de
Maestría en Ciencias de la Computación

Director de tesis
Dr. René Santaolaya Salgado

Cuernavaca, Morelos, México. Enero de 2020.



"2019, Año del Caudillo del Sur, Emiliano Zapata"

Cuernavaca, Mor.,
Oficio No.
Asunto:

26/noviembre/2019
DCC/088/2019
Aceptación de
documento de tesis
CENIDET-AC-004-M14

DR. GERARDO VICENTE GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de la **C. Ing. Heidi Barrera Monje**, con número de control **M17CE085**, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado "**Evaluación Experimental para Determinar la Relación entre las Métricas de Cohesión y Acoplamiento en Arquitecturas Orientadas a Objetos**" y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

Dr. René Santaolaya Salgado
Director de tesis

Dra. Olivia Graciela Fragoso Díaz
Revisor 1

M.C. Mario Guillén Rodríguez
Revisor 2

C.p. M.E. Guadalupe Garrido Rivera - Jefa del Departamento de Servicios Escolares.
Estudiante
Expediente

JGGS/lmz



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico
Subdirección Académica

"2019, Año del Caudillo del Sur, Emiliano Zapata"

Cuernavaca, Mor.,
No. de Oficio:
Asunto:

12/diciembre/2019
SAC/360/2019
Autorización de
impresión de Tesis

ING. HEIDI BARRERA MONJE
CANDIDATO AL GRADO DE MAESTRO EN CIENCIAS
DE LA COMPUTACIÓN
PRESENTE

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado "Evaluación Experimental para Determinar la Relación entre las Métricas de Cohesión y Acoplamiento en Arquitecturas Orientadas a Objetos", ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.



SEP TecNM
CENTRO NACIONAL
DE INVESTIGACIÓN
Y DESARROLLO
TECNOLÓGICO
SUBDIRECCIÓN
ACADÉMICA

ATENTAMENTE

Excelencia en Educación Tecnológica
"Conocimiento y tecnología al servicio de México"

DR. GERARDO VICENTE GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO

C.p. M.E. Guadalupe Garrido Rivera.- Jefa del Departamento de Servicios Escolares.
Expediente

CVGR/ego

Interior Internado Palmira S/N, Col. Palmira, C. P. 62490, Cuernavaca, Morelos.
Tel. (01) 777 3 62 77 70, ext. 4104, e-mail: acad_cenidet@tecnm.mx

www.tecnm.mx | www.cenidet.edu.mx



Agradecimientos

A Dios por la fortaleza que le ha dado a mi corazón porque a pesar de las circunstancias se culminó con éxito este trabajo de tesis.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) y al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), por el apoyo brindado para la realización del presente trabajo de tesis.

Al Dr. René Santaolaya Salgado director de esta tesis, por su aportación y supervisión continúa en el desarrollo de esta investigación.

A los revisores la Dra. Olivia Graciela Fragoso Díaz por las observaciones y comentarios los cuales ayudaron a realizar un mejor trabajo de investigación.

Y al M.C. Mario Guillen Rodríguez por el tiempo dedicado en la revisión del presente documento.

A mis compañeros y amigos de generación, por compartir sus conocimientos y por el apoyo brindado en todo momento.

Por último, pero no menos importante a toda mi familia por su apoyo moral y emocional. A mis hijos Gael y Aitana, por su paciencia y amor, y a mi esposo Ricardo por su apoyo incondicional para concluir este trabajo de investigación.

Dedicatorias

A Dios que me guía y me cuida en cada paso que doy y me da fortaleza para continuar.

A mis hijos Aitana y Gael por su amor incondicional, la alegría y fuerza que le dan a mi vida.

A mi esposo Ricardo por su apoyo y motivación, porque este logro es compartido con él.

A mis padres: Esperanza y Honorio que me han educado con principios y valores y siempre se han preocupado por mi bienestar.

A mis hermanas por motivarme a seguir adelante y a concluir con éxito todas las metas.

Resumen

Uno de los principales objetivos del análisis y diseño orientado a objetos es implementar sistemas de software, que se facilite su mantenimiento y sus componentes sean reusables. Cuando un componente o unidad de programa es autónomo, significa que no necesita de otros componentes para cumplir con su objetivo o meta de valor, es decir tiene independencia funcional, facilitando su reuso y su mantenimiento. La autonomía está en función de los atributos de cohesión y acoplamiento. Se dice que un buen diseño arquitectural, debe mantener en equilibrio el nivel de cohesión y el nivel de acoplamiento (alta cohesión y un bajo acoplamiento), en las unidades de programa que participan colaborativamente para satisfacer requerimientos.

El problema es que aún no se tiene conocimiento acerca del sentido y grado de influencia que tienen entre sí estos atributos. En esta tesis se realizó una evaluación experimental para aportar conocimiento acerca del comportamiento de la medida de acoplamiento entre unidades de programa con diferentes niveles de cohesión.

La evaluación inició con el desarrollo de un diseño experimental, donde se definieron las hipótesis, se identificaron las variables, se diseñó el plan de pruebas y se realizó la experimentación pertinente sobre el fenómeno de estudio. Para las pruebas experimentales, se obtuvieron las mediciones de cohesión y acoplamiento a nivel de clase y a nivel de sistema de 5 casos de prueba, que fueron re-factorizados en 12 arquitecturas diferentes cada uno. Para esta fase, se desarrolló un marco orientado a objetos para realizar las mediciones de forma automática.

Para la interpretación de resultados del estudio experimental, se utilizó el enfoque de análisis de regresión y de correlación. Este análisis permitió conocer el sentido y grado de relación de las métricas de cohesión y de acoplamiento, el cual resultó ser inversamente proporcional, es decir cuando una aumenta, la otra disminuye. También se encontró la función que mejor representa la relación entre estas métricas. Dando una función polinómica de segundo grado.

Palabras clave: métrica, cohesión, acoplamiento, análisis de regresión, correlación, función.

Abstract

One of the main objectives of object-oriented analysis and design is to implement software systems, which facilitate its maintenance and its components are reusable. When a component or program unit is autonomous, it means that it does not need other components to fulfill its objective or value goal, that is to say, it has functional independence, facilitating its reuse and maintenance. Autonomy is depending of cohesion and coupling attributes. A good architectural design, must keep in balance the level of cohesion and the level of coupling (high cohesion and low coupling) in the program units that participate collaboratively to meet requirements.

The problem is that no has knowledge about the meaning and degree of influence that these attributes have on each other. In this thesis, an experimental study that provides knowledge about the behavior of the coupling measure between program units with different levels of cohesion was conducted.

The experimental study began with the development of an experimental design, where the hypotheses were defined, the variables were identified, the test plan was designed and the relevant experimentation was carried out on the study phenomenon. For the experimental tests, cohesion and coupling measurements at the class level and at the system level for 5 test cases were obtained, which were re-factored into 12 different architectures each. For this phase, an object-oriented framework to perform measurements automatically was developed.

For the interpretation of results about the experimental study, the regression and correlation analysis approach was used. This analysis allowed to know the meaning and relationship degree of the cohesion and coupling metrics, which is inversely proportional, that is when one increases the other decreases. The function that best represent the relationship between these metrics was also found. Giving a second degree polynomial function.

Keywords: metrics relationship, cohesion, coupling, regression analysis, correlation, math function

Lista de figuras.....	iii
Lista de tablas.....	iv
Glosario de términos.....	v
Capítulo 1: Introducción	1
1.1 Introducción	1
1.2 Organización del documento de tesis.....	2
Capítulo 2: Marco de Referencia	4
2.1 Trabajos relacionados.....	4
2.2 Trabajos antecedentes en el CENIDET.....	7
2.3 Objetivos de la tesis.....	8
2.4 Justificación.....	8
2.5 Planteamiento del problema	9
2.6 Alcances y limitaciones de la investigación.....	9
Capítulo 3: Marco Teórico.....	10
3.1 Introducción al Diseño experimental	10
3.2 Calidad del software.....	11
3.3 Métricas de diseño.....	11
3.3.1. Métricas de Cohesión.....	12
3.3.2. Métricas de Acoplamiento	14
3.4 Análisis de regresión y de correlación	16
3.4.1. Tipos de regresión.....	17
3.4.2. Método de mínimos cuadrados	18
3.4.3. Coeficiente de correlación	20
3.4.4. Coeficiente de determinación	21
Capítulo 4: Diseño Experimental	22
4.1 Formulación de preguntas e hipótesis	22
4.1.1. Pregunta de investigación	22
4.1.2. Formulación de hipótesis	22
4.2 Identificación de variables	23
4.2.1 Variables independientes.....	23

4.2.2 Variables dependientes	23
4.3 Diseño del plan de pruebas.....	24
4.3.1 Descripción de los casos de prueba	24
4.3.2 Diseño de los casos de prueba	26
4.3.3 Desarrollo del instrumento de pruebas	34
Capítulo 5: Ejecución del Experimento y Recolección de Mediciones.....	36
5.1 Ejecución del experimento	36
5.2 Recolección de mediciones	37
Capítulo 6: Análisis Estadísticos e Interpretación de Resultados	45
6.1 Análisis estadístico.....	45
6.1.1 Análisis #1: Mediciones de cohesión y acoplamiento a nivel de clase	46
6.1.2 Análisis #2: Mediciones de cohesión y acoplamiento a nivel de sistema	56
Capítulo 7: Conclusiones y Trabajos Futuros.....	66
7.1 Conclusiones	66
7.2 Trabajos futuros.....	67
Referencias Bibliográficas.....	68

Lista de figuras

Figura 1. Arquitectura del escenario#1.....	27
Figura 2. Arquitectura del escenario#2.....	27
Figura 3. Arquitectura del escenario#3.....	28
Figura 4. Arquitectura del escenario#4.....	28
Figura 5. Arquitectura del escenario#5.....	29
Figura 6. Arquitectura del escenario#6.....	29
Figura 7. Arquitectura del escenario#7.....	30
Figura 8. Arquitectura del escenario#8.....	30
Figura 9. Arquitectura del escenario#9.....	31
Figura 10. Arquitectura del escenario#10.....	32
Figura 11. Arquitectura del escenario#11.....	33
Figura 12. Arquitectura del escenario#12.....	34
Figura 13. Arquitectura de clase del Marco de métricas orientado a objetos.....	35
Figura 14. Mediciones a nivel de clase del CP. Bernoulli.....	46
Figura 15. Mediciones a nivel de clase del CP. Continuas.....	48
Figura 16. Mediciones a nivel de clase del CP. Discretas.....	49
Figura 17. Mediciones a nivel de clase del CP. Hipercubo.....	50
Figura 18. Mediciones a nivel de clase del CP. TaylorBucket.....	51
Figura 19. Función polinómica y r^2 de cohesión y CoupT – CP. TaylorBucket.....	53
Figura 20. Función polinómica y r^2 de cohesión y CoupT – CP.Hipercubo.....	55
Figura 21. Función polinómica y r^2 de cohesión y CoupT – CP. Continuas.....	56
Figura 22. Función polinómica y r^2 de cohesión y COF – CP.Bernoulli.....	58
Figura 23. Función polinómica y r^2 de cohesión y WTCoup – CP. Bernoulli.....	58
Figura 24. Función lineal y r^2 de cohesión y COF – CP. Continuas.....	60
Figura 25. Función lineal y r^2 de cohesión y WTCoup – CP. Continuas.....	60
Figura 26. Función lineal y r^2 de cohesión y COF – CP. Discretas.....	61
Figura 27. Función lineal y r^2 de cohesión y WTCoup – CP. Discretas.....	62
Figura 28. Función polinómica y r^2 de cohesión y COF – CP. Hipercubo.....	63
Figura 29. Función polinómica y r^2 de cohesión y WTCoup – CP. Hipercubo.....	63
Figura 30. Función polinómica y r^2 de cohesión y COF – CP. TylorBucket.....	64
Figura 31. Función polinómica y r^2 de cohesión y WTCoup – CP. TylorBucket.....	65

Lista de tablas

Tabla 1. Tabla comparativa de trabajos relacionados.....	6
Tabla 2. Valores predefinidos para los casos que no cumplen con la restricción.	36
Tabla 3. Mediciones a nivel de clase CP. Bernoulli.....	37
Tabla 4. Mediciones CP. V.A.Continuas	38
Tabla 5. Mediciones CP. V.A. Discretas.....	39
Tabla 6. Mediciones CP. Hipercubo.....	40
Tabla 7. Mediciones CP. TaylorBucket	41
Tabla 8. Mediciones a nivel de sistema CP. Bernoulli.....	42
Tabla 9. Mediciones del CP. Continuas	42
Tabla 10. Mediciones del CP. Discretas.....	43
Tabla 11. Mediciones del CP. Hipercubo.....	43
Tabla 12. Mediciones del CP. TaylorBucket.....	44
Tabla 13. Sumatorias de LCOM* y CoupT.....	47
Tabla 14. Indicador de r para cohesión y CoupT – CP. Bernoulli	47
Tabla 15. Indicador de r para cohesión y CoupT - CP. Continuas.....	48
Tabla 16. Indicador de r para cohesión y CoupT - CP. Discretas	49
Tabla 17. Indicador de r para cohesión y CoupT - CP. Hipercubo	50
Tabla 18. Sumatorias de Coh y CoupT	52
Tabla 19. Sumatorias de los errores al cuadrado - Coh y CoupT.....	54
Tabla 20. Indicador de r para cohesión y CoupT - CP. TaylorBucket.....	54
Tabla 21. Indicador de r para cohesión y CoupT - CP. Hipercubo	55
Tabla 22. Indicador de r para cohesión y CoupT - CP. Continuas.....	56
Tabla 23. Indicador de r para cohesión con acoplamiento - CP. Bernoulli.....	57
Tabla 24. Indicador de r para cohesión con acoplamiento - CP. Continuas.....	59
Tabla 25. Indicador de r de cohesión con acoplamiento - CP. Discretas	61
Tabla 26. Indicador de r de cohesión con acoplamiento - CP. Hipercubo.....	62
Tabla 27. Indicador de r de cohesión con acoplamiento - CP. TylorBucket.....	64

Glosario de términos

Cohesión	Es un indicador cualitativo del grado en el que un módulo se centra en hacer una sola cosa (Roger S. Pressman, 2010).
Acoplamiento	Es una medida de la fuerza de asociación establecida por una conexión de un módulo a otro (Yadav, Sikka, & Shrivastava, 2014).
Reusabilidad	Es la habilidad de los elementos de software para servir en la construcción de muchas aplicaciones diferentes (Strub, 2015).
Modularidad	Es la propiedad que permite dividir una aplicación en partes más pequeñas (llamadas módulos), cada una de las cuales debe ser tan independiente como sea posible de la aplicación en sí y de las partes restantes.(Liskov, 1980).
Encapsulamiento	Es el empaquetamiento (o enlazado) de una colección de elementos. El encapsulamiento comprende las responsabilidades de una clase, incluyendo sus atributos (y otras clases para objetos agregados) y operaciones, así como los estados de la clase, según se definen mediante valores específicos de atributos (Sánchez, 2010).
Meta de valor	En el contexto de esta investigación una meta valor u objetivo, es relativa a la unidad de programa (función, clase, modulo o sistema). A nivel de función, significa realizar una única operación funcional. A nivel de clases, modelar el comportamiento de una entidad. A nivel de módulo o subsistema, es atender un requerimiento específico. A nivel de sistema, una meta de valor es resolver una aplicación completa o proceso de negocio planteada por un cliente o usuario.
Métrica	Es el proceso por el cual se obtiene una medida cuantitativa del grado en que un sistema, componente o proceso posee un atributo dado (IEEE, 1993).
Medida	Es el valor asignado a un atributo de una entidad mediante una medición (IEEE, 1993).

Medición Es el acto de determinar una medida (Roger S. Pressman, 2010).

Modelo matemático Es una construcción matemática abstracta y simplificada, relacionada con una parte de la realidad y creada para un propósito particular. Así, por ejemplo, un gráfico, una función o una ecuación, pueden ser modelos matemáticos de una situación específica (Dirac, 1984).

Capítulo 1: Introducción

1.1 Introducción

En la actualidad el enfoque orientado a objetos ha alcanzado altas cuotas de popularidad debido a los beneficios que promete: modularidad, reusabilidad, extensibilidad, etc., características que inciden directamente en la mejora de la calidad de los productos de software. Se piensa que los programas construidos utilizando técnicas de programación orientada a objetos (POO) automáticamente cuentan con estas características. Esto es falso, debido a que estos beneficios dependen de la habilidad y la experiencia del diseñador del software. Cuando el desarrollador de Software carece de estas habilidades, suelen producirse unidades de programa que resultan poco reusables y costosas en su mantenimiento, debido a que los sistemas de software exhiben incoherencia en sus partes o módulos, provocando que existan problemas de insuficiencia, completitud y dependencias, cuando se trata de alcanzar su objetivo o meta de valor.

Las dependencias entre módulos de programa se manifiestan debido a la carencia de un efectivo encapsulamiento. El encapsulamiento está en función del grado de relación que tienen las funciones dentro de los módulos o unidades de software (tales como clases), así mismo entre clases o módulos. Esta relación se da por la interacción entre las funciones (esto es la capacidad funcional) y por los datos accedidos por estas funciones.

Las relaciones funcionales se presentan de manera directa o de manera indirecta. Esto es cuando explícitamente se comparten datos o responsabilidades funcionales de manera directa. El otro caso se da cuando se comparte información de manera indirecta por medio del acceso a variables de mayor alcance, como lo pueden ser las variables o funciones globales, públicas y parcialmente compartidas.

El concepto de relación entre clases es una medida relativa de la interacción entre éstas y la cantidad de información que fluye entre ellas. Es relativa, porque para atender un objetivo o meta de valor éste puede ser a nivel de función, de clase, de requerimiento o de un subsistema o sistema completo y depende de la capacidad funcional requerida en esos niveles de abstracción.

Cuando un módulo no es suficientemente cohesivo, puede faltarle capacidad funcional o que no cuente con los datos suficientes para alcanzar el objetivo o meta de valor requerido (esto quiere decir que no es suficiente o completo para alcanzar una meta u objetivo). Es cuando la unidad de programa (ya sea una función, clase, módulo o sistema) necesita de la interacción con otras unidades de programa, tanto para acceder a datos o para extender su capacidad

funcional. Esto incrementa las dependencias y el acoplamiento entre las unidades de programa involucradas.

Por lo tanto, la cohesión y el acoplamiento son dos atributos importantes que denotan la calidad a nivel de diseño de un sistema de software. Son un indicador importante de la modularidad que favorece el encapsulamiento y ayudan a crear unidades autónomas de software que son directamente reutilizables debido a su menor dependencia con otras unidades.

Un buen diseño arquitectural, debe mantener en equilibrio el nivel de cohesión y el nivel de acoplamiento (alta cohesión y un bajo acoplamiento), para lograr beneficios en costos por mantenimiento y reuso de las entidades de programa en sus diferentes niveles de granulación y atención de responsabilidades. El problema radica en que no se tiene conocimiento acerca del sentido y grado de influencia que tienen entre sí estas métricas de calidad de software, por lo que no se sabe cuáles son las dimensiones correctas entre estas dos métricas para alcanzar ese equilibrio.

En este trabajo de investigación, se realizó un estudio experimental que proporciona conocimiento acerca del sentido y grado de influencia que tienen entre sí las métricas de calidad de software: cohesión y acoplamiento, de diferentes unidades de programa que participan colaborativamente para satisfacer requerimientos.

1.2 Organización del documento de tesis

El contenido del presente documento de tesis es el siguiente:

El capítulo dos “Marco de referencia” describe el estado del arte, los trabajos relacionados y los trabajos de antecedente realizados en el CENIDET. Posteriormente, se encuentran la descripción de los objetivos, el planteamiento del problema y los alcances y las limitaciones del presente trabajo de tesis.

El capítulo tres “Marco teórico”, incluye la teoría de las métricas de cohesión y acoplamiento, los métodos y los procedimientos para el análisis estadístico, requeridos en este trabajo de investigación.

El capítulo cuatro “Diseño experimental”, describe detalladamente las etapas del diseño experimental, la pregunta de investigación, la formulación de las hipótesis, las variables que intervienen en el experimento y el diseño del plan de pruebas; donde se especifica y se describen los escenarios de los casos de prueba que se usaron para la medición de las

métricas. Así como la herramienta utilizada como instrumento de prueba para la ejecución del experimento.

El capítulo cinco “Ejecución del experimento y recolección de mediciones”, explica cómo se llevaron a cabo las pruebas, su clasificación y se muestran todas las mediciones obtenidas.

El capítulo seis “Análisis estadísticos e interpretación de resultados”, muestra el análisis realizado para la interpretación de las mediciones, se describe cada uno de los métodos y se concluye con el modelo matemático resultante del análisis.

El trabajo de tesis concluye con el capítulo siete “Conclusiones y trabajos futuros”, en donde se discuten los resultados obtenidos, se establecen las conclusiones finales y se proponen los trabajos futuros a desarrollar.

Capítulo 2: Marco de Referencia

2.1 Trabajos relacionados

La búsqueda de trabajos relacionados con esta tesis, se hizo en revistas indexadas con un factor de impacto de 0.4 a 1.0 y fecha de publicación dentro de los últimos 6 años (2013-2019). Algunas de las cadenas de búsqueda utilizadas son: “*Relationship between cohesion and coupling*”, “*Metric that define the cohesion and coupling behavior in object-oriented software*”, “*Relationship between static quality (QoS) metrics cohesion and its coupling factor*”, entre otras.

Para evaluar los trabajos encontrados se realizó un análisis con las siguientes fases o criterios:

Lo que se aclama. Se considera este criterio para conocer los beneficios que se obtienen en cada uno de los trabajos, en comparativa con este trabajo de tesis. Algunos autores aclaman que mejoran la cohesión de un sistema de software, otros aclaman un valor de umbral para cada tipo de métrica de acoplamiento, otros que evalúan la confiabilidad y la mantenibilidad del software mediante las métricas de cohesión y acoplamiento, y muchos más que encuentran la correlación entre algunas métricas, como “entre mayor número de clases mayor es el valor de acoplamiento”. Ninguno de los trabajos analizados aclama un valor que represente el grado de cambio de una métrica con respecto a otra. Esto es lo que se presenta en esta tesis.

Objetivo. Se considera este criterio de comparación para establecer la diferencia en cuanto al objetivo que se persigue con la aportación de cada uno de los trabajos estudiados, en comparativa con el objetivo planteado en esta tesis. El objetivo de algunos de los trabajos analizados, es la creación de alguna nueva métrica para mejorar la cohesión, mejorar la precisión en el cálculo de las ya existentes o bien, cómo influye la cohesión y el acoplamiento en los atributos de calidad externa del software. Ninguno de los trabajos analizados tiene como objetivo encontrar la relación existente entre estas métricas de calidad y representar dicha relación de manera formal.

Método. Los trabajos analizados difieren en el enfoque de solución para alcanzar su objetivo. El enfoque de patrones de uso frecuente (FUP) que parte de la frecuencia de interacción de las variables entre los módulos. El enfoque de lógica difusa (LD), que proporciona un mecanismo de inferencia, el cual permite simular los procedimientos de razonamiento humano en sistemas basados en el conocimiento, facilitando el modelado de información cualitativa de forma aproximada. El enfoque de regresión logística univariante y multivariante (RLU-M), este método explica o predice el valor de una variable en función de sus variables independientes. Y por lo último, tenemos el enfoque experimental (EE), que

parte de controlar deliberadamente las variables para delimitar relaciones entre ellas, este enfoque está basado en la metodología científica.

Métricas utilizadas. Este criterio indica si los trabajos analizados se basaron o utilizaron alguna de las métricas existentes en la literatura. Por ejemplo: Carencia de cohesión (LCOM), Acoplamiento (CBO), Número de clases (NCLA), Tamaño (NOM Numero de métodos, NOA Numero de atributos, LOC Líneas de código), Número de canales de comunicación (NCC), etc.

Aportación. Este criterio se utiliza para establecer el grado de innovación u originalidad de los diferentes trabajos relacionados, se consideran trabajos que ofrecen una nueva Métrica (MTR), Algoritmo (ALG), Método (MET), Herramienta (HER), Análisis (ANL), Modelo (MOD), o alguna combinación de éstas. Se considera que la aportación tiene directamente un impacto total o parcial, en el alcance del objetivo de cada enfoque estudiado.

Alcance. Este criterio se utiliza para definir el alcance que tiene cada aportación en los trabajos analizados, éste puede ser: Identificado (IND), Implementado (IMP) y/o Probado (PROB).

Producto o Resultado. Este es el producto derivado de la aportación realizada, el cual sirve de instrumento para alcanzar, parcial o totalmente, el objetivo del trabajo. En este contexto, hay trabajos relacionados que aportan como producto o resultado: alguna nueva métrica o la modificación a alguna existente tal como Cohesión, Acoplamiento, Coherencia, entre otras. Así como: herramientas para la medición de las métricas, algún método, modelo o simplemente un análisis.

Tabla 1. Tabla comparativa de trabajos relacionados

Autor	Lo que aclama	Objetivo	Método	Métrica usada	Aporte	Alcance	Resultado
(Rathee & Chhabra, 2018)	Mejorar la cohesión y disminuir el acoplamiento simultáneamente	Proponer una nueva métrica de cohesión, basada en Patrones de Uso Frecuente a nivel de módulo.	Métrica y algoritmo de agrupación basado en patrones de uso frecuente (FUP).	LCOM1 LCOM2 LCOM3	MTR ALG	IMP PROB	UPBC, FUPClust
(Kalantari, Alizadeh & Motameni, 2015)	Evaluación de la confiabilidad del software mediante cohesión y acoplamiento	Evaluar la confiabilidad de software obteniendo mayor precisión en el cálculo de los índices de cohesión y acoplamiento.	Tipos de acoplamiento y cohesión de acuerdo al modelo de Myers. Calculo total usando técnicas de lógica difusa.	LCOM CBO	MTR MET ALG	MTR IMP MET IDN	<i>cohesion coupling</i>
(Bivde & Sarasu, 2016)	Entre mayor número de clases mayor es el valor de acoplamiento.	Encontrar la correlación entre los valores de métricas de acoplamiento y el número de clases en proyectos multimedia de Java.	Experimental	CBO NCLA	HER	IMP PROB	JCMT ANL
(Bidve & Sarasu, 2016)	Un valor de umbral para cada tipo de métrica de acoplamiento	Encontrar el impacto de cada métrica de acoplamiento en varios atributos de calidad de software.	Experimental	CBO	ANL	IMP	Análisis
(Al Dallal, 2013)	La mantenibilidad de la clase mejora al reducir su tamaño y el acoplamiento y al aumentar su cohesión.	Encontrar la relación entre los atributos de calidad interna de la clase y un atributo de calidad externa (mantenibilidad de la clase)	La regresión logística univariante y multivariante	LCOM CBO Size	MOD	IMP	Modelo univariante y multivariante
Este trabajo de tesis	Sentido, grado y forma de relación entre las métricas de cohesión y acoplamiento mediante un modelo matemático.	Obtener conocimiento acerca de la relación existente entre las métricas de calidad estática. Cohesión y Acoplamiento.	Experimental	LCOM*, LCOM5, ClassCo y Coh, WTCoup y COF	HER MOD	IMP PROB	Modelo matemático

En la tabla 1 se muestra una comparativa entre los principales trabajos relacionados con esta tesis, utilizando las fases o criterios definidos anteriormente. Los trabajos analizados difieren en su objetivo con este trabajo de tesis. En el trabajo (Rathee & Chhabra, 2018), los autores muestran cómo se mejora la cohesión en un sistema de software. En el trabajo (Kalantari et al., 2015), los autores evalúan la confiabilidad del software mediante cohesión y acoplamiento. En (Bivde & Sarasu, 2016), se busca la correlación entre los valores de métricas de acoplamiento y el número de clases. En el trabajo (Bidve & Sarasu, 2016), se busca el impacto de cada tipo de acoplamiento en los atributos de calidad del software. Por último, en (Al Dallal, 2013) se busca la relación entre las métricas de tamaño y cohesión y su impacto en el atributo de mantenibilidad de la clase.

En contraste a los trabajos descritos anteriormente, en este trabajo de tesis, se describe la metodología para conocer el sentido y grado de relación entre las métricas de cohesión y acoplamiento, así como la función matemática que mejor representa dicha relación. Ninguno de los trabajos mencionados tiene el mismo objetivo de esta tesis, sin embargo, se utilizaron las métricas de cohesión y acoplamiento propuestas en el trabajo (Kalantari et al., 2015), debido a su mayor precisión en su cálculo. Adicionalmente, a diferencia del trabajo (Al Dallal, 2013), en que el análisis de regresión logística univariante y multivariante considera que la relación entre métricas es lineal, en este trabajo de tesis se determinó que la relación entre las métricas de cohesión y acoplamiento no es lineal, sino polinómica de segundo grado.

La ventaja de este trabajo de tesis en relación con los trabajos analizados, es que no sólo se llevó a cabo un método experimental para conocer la relación entre las métricas, sino además incluye una función matemática que representa formalmente el grado de variación de la métrica de acoplamiento con respecto a la métrica de cohesión.

2.2 Trabajos antecedentes en el CENIDET

El presente trabajo de tesis está relacionado con el siguiente trabajo realizado anteriormente en el CENIDET:

Estudio Experimental para Determinar la Influencia del Nivel de Granulación de Servicios Web en Factores de Calidad QoS (Rogel, 2014). Este trabajo es de investigación teórica, que trata de dar respuesta a la pregunta *¿Cuál es el nivel adecuado de granulación de servicios web?*, mediante un estudio experimental completo cuyo objetivo es obtener conocimiento acerca de cómo influyen ciertos niveles de granulación de servicios Web sobre el tiempo de respuesta. Así mismo, se desarrolló una herramienta de software que sirvió como instrumento de pruebas para recopilar la información para el análisis correspondiente al estudio experimental.

2.3 Objetivos de la tesis

- **General**

Obtener conocimiento acerca del sentido y grado de relación existente entre las métricas de calidad estática Cohesión y Acoplamiento.

- **Específico**

Formalizar la relación de estas métricas mediante un modelo matemático.

2.4 Justificación

En la POO, unas de las propiedades que deben tener los módulos para ser considerados auténticos objetos, es el encapsulamiento y su reuso. Adicionalmente, pueden ser consideradas otras propiedades tales como: alta cohesión (funcionalidad bien delimitada) y bajo acoplamiento (interdependencia funcional limitada). La presencia de estas características en la proporción adecuada, determina de forma significativa la calidad del software.

Una unidad de programa muy grande que exhibe incorrectos niveles de cohesión, podría contar con funcionalidad sobrada que a lo mejor no se utiliza para alcanzar una meta de valor, afectando su legibilidad y movilidad. Por otro lado, una unidad de programa muy pequeña podría ser incompleta para satisfacer una meta de valor, por lo tanto, no sería suficiente y completa, de manera que necesitaría de otras unidades de programa para completar las funcionalidades necesarias para satisfacer los requerimientos de software; lo cual incrementaría su dependencia y su factor de acoplamiento entre las unidades de programa participantes. Esta situación afecta la continuidad y la protección modular, así como la movilidad para efectos de mantenimiento y reuso.

(Beck & Diehl, 2011), proponen que las características a considerar para una buena modularidad es que cada módulo debe tener alta cohesión y un bajo a acoplamiento.

En esta tesis se considera que un buen diseño arquitectural, debe mantener en equilibrio los factores de calidad de cohesión y acoplamiento, logrando beneficios en costos por mantenimiento y reuso. El estudio realizado que se presenta en este documento, aporta conocimiento acerca del sentido y grado de influencia que tienen entre sí estos factores, así como el comportamiento de las unidades de programa en cuanto a su acoplamiento, cuando se varía la cohesión en éstas.

2.5 Planteamiento del problema

El problema radica en que algunas unidades de programa exhiben alguna o varias de las siguientes características: inmovilidad, rigidez, insuficiencia, lo cual impide la reusabilidad. Una causa importante que origina este problema es que no se tiene conocimiento acerca del sentido y grado de influencia que tienen entre sí las métricas de calidad de software. Específicamente, no se tiene conocimiento del grado de relación que tienen las métricas de cohesión y acoplamiento, entre diferentes unidades de programa que participan colaborativamente para satisfacer requerimientos.

2.6 Alcances y limitaciones de la investigación

Alcance

- Este trabajo de tesis se enfocó en los factores de cohesión y acoplamiento, a nivel de clases y a nivel de sistema.
- Se determinó el grado de variación de la métrica de acoplamiento en relación al cambio de la métrica de cohesión.
- Se realizó el análisis estadístico para comprobar la veracidad de los resultados obtenidos durante el periodo de pruebas.
- Se representó la relación de las métricas a través de un modelo matemático.

Limitaciones

- Sólo fue considerada la relación entre las métricas de cohesión y acoplamiento descritas en la sección 3.3 de este documento.
- Para las pruebas solamente se consideraron los casos de prueba referidos en el trabajo de antecedente (Rogel, 2014).
- Se utilizaron los métodos de regresión y correlación para el análisis de los datos.

Capítulo 3: Marco Teórico

3.1 Introducción al Diseño experimental

(Fisher, 1935), el padre del diseño experimental moderno, observó que los experimentos son *“solo experiencia cuidadosamente planificada de antemano y diseñados para formar una base segura de nuevos conocimientos”*.

El diseño de experimentos es la aplicación del método científico para generar conocimiento acerca de un sistema o proceso, por medio de pruebas planeadas adecuadamente. Esta metodología se ha ido consolidando como un conjunto de técnicas estadísticas y de ingeniería, que permiten entender mejor, situaciones complejas de relación causa-efecto (Pulido & De la Vara, 2008).

En un diseño experimental se manipulan deliberadamente una o más variables independientes para observar y analizar los efectos que tiene sobre una o más variables dependientes, en una situación controlada.

Para efectos de este trabajo de investigación se definen cinco fases para el estudio experimental:

1. Formulación de la(s) pregunta(s) de investigación y las hipótesis. Se plantean las preguntas con respecto a los objetivos de la investigación. Se definen las hipótesis general y alternativa, en términos de las variables del fenómeno a examinar. Es importante definir un procedimiento cuantitativo (hipótesis estadísticas) con el fin de evaluar las hipótesis.
2. Identificación de variables. Se seleccionan las variables que mejor reflejen el fenómeno de estudio; variables independientes, dependiente e intervinientes.
3. Diseño del plan de pruebas. Se establece el número de muestras a recolectar y se definen los sujetos de prueba sobre los cuales se obtendrán las muestras.
4. Ejecución del experimento. Se ejecuta el experimento de acuerdo al diseño, para recolectar los datos que serán analizados. Entre mayor sea la cantidad de datos recolectados se tendrá un mejor punto de comparación, para establecer la relación entre las variables.
5. Análisis de resultados. Se analizan los datos obtenidos durante el experimento, para validar las hipótesis planteadas y encontrar las relaciones entre las variables. En esta

6. fase, se aplican conceptos estadísticos para analizar los datos e identificar correlaciones o relaciones entre las variables.

3.2 Calidad del software

La calidad del software es *“el grado con el que un sistema, componente o proceso cumple los requerimientos especificados y las necesidades o expectativas del cliente o usuario”* (IEEE, 1993). La calidad del software es el conjunto de cualidades que lo caracterizan y que determinan su utilidad y existencia. La mayoría de los modelos de calidad existentes están basados en la norma ISO 9126:2001. Esta norma define un conjunto de características de calidad que son refinadas en sub características que están descompuestas en atributos. La calidad del software es medible y se hace en base a sus atributos. La medida de estos atributos se calcula mediante la utilización de métricas.

Las métricas las podemos ver desde una perspectiva tridimensional, con las siguientes dimensiones (Escudero, 2001):

- Características o atributos internos del software a medir (por ejemplo, cohesión, acoplamiento, complejidad...).
- Etapa del ciclo de vida en la que se puede medir (análisis, diseño, codificación, implementación, pruebas...).
- Nivel de granulación en el que se mide (por ejemplo, nivel de sistema, de programa, de clase, de método o de variable).

La presente tesis se enfoca en las métricas conocidas en la literatura como métricas de diseño, que miden los atributos de cohesión, canales de comunicación y acoplamiento a nivel de clase y a nivel de sistema.

3.3 Métricas de diseño

Uno de los principales objetivos del análisis y diseño orientado a objetos es implementar un sistema de software donde los componentes tengan una alta cohesión y un bajo acoplamiento, facilitando su reuso (Paixao, Harman, Zhang, & Yu, 2017).

La cohesión y el acoplamiento son dos atributos internos que denotan la calidad del diseño de un sistema de software. Permite al ingeniero de software evaluar la calidad durante el desarrollo del sistema. Estos atributos influyen en los factores externos de calidad del software como la confiabilidad, la eficiencia, la integridad, el mantenimiento, la flexibilidad, la portabilidad, su reuso y la interoperabilidad del software orientado a objetos (Bidve & Sarasu, 2016). Son un indicador importante de la modularidad que favorece la encapsulación

y ayudan a crear componentes de software que son directamente reusables debido a su mayor autonomía con otros componentes (Rathee & Chhabra, 2018).

3.3.1. Métricas de Cohesión

“La cohesión es un indicador cualitativo del grado en el que un módulo se centra en hacer una sola cosa” (Roger S. Pressman, 2010).

El término cohesión tiene su origen en el diseño estructural y se refiere a la relación entre los diversos elementos de un módulo dado. Es un indicador importante de la calidad del diseño del software y de su modularidad. Un valor de cohesión alto en un módulo indica que éste proporciona una funcionalidad casi única, mientras que un valor bajo indica que el módulo tiene más de una responsabilidad funcional. Entonces, un módulo con mayor cohesión es siempre deseable, ya que es más comprensible, modificable, mantenible y reusable (Rathee & Chhabra, 2018)

Para efectos de la presente tesis se analizaron un conjunto de métricas de cohesión y se consideraron las siguientes:

- **LCOM* Carencia de Cohesión en los métodos (Henderson-Sellers, 1996)**

Esta métrica considera un grupo de métodos $\{Mi\}$ ($i=1, \dots, m$) accediendo a un conjunto de atributos $\{Aj\}$ ($j=1, \dots, a$) y el número de métodos que acceden a cada atributo como $\mu(Aj)$.

$$LCOM^* = \frac{\left(\frac{1}{a} \sum_{j=1}^a \mu(A_j)\right) - m}{1 - m} \quad (1)$$

Donde:

a = Número de atributos

$\mu(Aj)$ = Número de métodos que acceden a cada atributo

m = Número de métodos

Esta métrica sólo puede calcularse cuando $m > 1$. Cuando todos los métodos acceden a todos los atributos, entonces $\sum \mu(Aj) = ma$, por lo tanto, $LCOM^* = 0$. Esto indica una cohesión perfecta. Valores cercanos a 0 indican que la mayoría de los métodos accede a la mayoría de los atributos. Por el contrario, cuando cada atributo es accedido solo una vez por alguno de los métodos, entonces $\sum \mu(Aj) = a$, por lo tanto, $LCOM^* = 1$, lo cual indica falta de cohesión.

- **LCOM4 Carencia de Cohesión en los métodos (Hitz & Montazeri, 1995)**

LCOM4 es el número de conjuntos de métodos conectados, tanto por compartir una variable de la clase o por invocarse entre ellos.

La métrica LCOM4 considera un grafo no dirigido G , donde los vértices son los métodos de una clase y existe una arista entre dos vértices si los métodos M_i y M_j *comparten al menos un atributo o si M_i invoca a M_j o viceversa*. El número de conjuntos que se formen en el grafo indica el valor de cohesión en la clase.

Cuando la medida de LCOM4 es igual a 1, la clase es cohesiva, cuando la medida es igual o mayor a 2, la clase tiene más de una responsabilidad y debe dividirse. Cuando LCOM4 es igual a 0, la clase no tiene métodos.

- **LCOM5 Carencia de Cohesión en los métodos (Henderson-Sellers, 1996)**

$$LCOM5 = \frac{(a - kl)}{(l - kl)} \quad (2)$$

Donde:

l = Número de atributos

k = Número de métodos

a = Número de métodos que acceden a cada atributo

Esta métrica tiene el mismo rango que la métrica LCOM*. Cuando $LCOM5 = 0$ indica buena cohesión y cuando $LCOM5 = 1$ significa una falta de cohesión.

- **ClassCoh Cohesión de la clase (G. Gui & Scott, 2006)**

ClassCoh proporciona la medida de cohesión de la clase, sumando las similitudes de todos los pares de métodos y dividido por el número total de pares (G. Gui & Scott, 2006). Su rango de valores va de 0 a 1.

$$ClassCoh = \frac{\sum_{i,j=1}^m sim(i,j)}{m^2 - m} \quad (3)$$

Donde:

m = Número de métodos

$sim(i,j)$ = Similitud de dos métodos, es decir tienen variables de instancia en común

$$sim(i, j) = \frac{|V_i \cap V_j|}{|V_i \cup V_j|} \quad (4)$$

Donde:

V = Variables de instancia

Por último, la cohesión transitiva ponderada a nivel del sistema de software $WTCoh$, se define como la cohesión promedio de todas las clases de las cuales está compuesto:

$$WTCoh = \frac{\sum_{j=1}^n ClassCoh_j}{n} \quad (5)$$

Donde:

n = Número de clases en el sistema

- **Coh Cohesión (Briand LC, Daly J, 1998)**

Coh es una medida de la cohesión de una clase, midiendo el número de atributos comunes usados por diferentes métodos, indicando la calidad de la abstracción hecha en la clase (Gui Gui & Scott, 2009). Su rango de valores es de 0 a 1.

$$Coh = \frac{a}{kl} \quad (6)$$

Donde:

a = Número de métodos que acceden a cada atributo

k = Número de métodos de la clase

l = Número de atributos de la clase

3.3.2. Métricas de Acoplamiento

“El acoplamiento es un indicador cualitativo del grado en el que un módulo está conectado con otros y con el mundo exterior” (Roger S. Pressman, 2010).

El acoplamiento se ve como una medida del incremento de la complejidad, reduciendo el encapsulamiento y su posible reuso. Así mismo, limita la comprensibilidad y mantenibilidad del sistema (Sicilia, 2009).

Para efectos de la presente tesis se consideran las siguientes métricas de acoplamiento:

- **CBO Acoplamiento entre objetos (Chidamber & Kemerer, 1994)**

El CBO de una clase es el número de clases a las cuales ésta se encuentra asociada. Se produce dependencia entre dos clases cuando una clase usa métodos o variables de la otra clase. Las clases relacionadas por herencia no se toman en cuenta.

Cuanto más acoplamiento se da en una clase, más difícil será su movilidad para reuso. Además, las clases con alto acoplamiento dificultan su comprensibilidad y hacen más difícil el mantenimiento. Las clases deberían de ser lo más autónomas que sea posible. Al reducir el acoplamiento se reduce la complejidad, se mejora la modularidad y se promueve la encapsulación (Rodríguez & Harrison, n.d.).

- **WTCoup Acoplamiento transitivo ponderado (G. Gui & Scott, 2006)**

El acoplamiento transitivo ponderado de un sistema se define de la siguiente manera:

$$WTCoup = \frac{\sum_{i,j=1}^m Coup(i,j)}{m^2 - m} \quad (7)$$

Donde:

m = Total de clases del sistema

$Coup(i,j)$ = Es el grado de acoplamiento directo e indirecto de la clase i con la clase j

Para efectos de este trabajo de tesis, solo se considera el acoplamiento directo entre la clase i y la clase j , definido como $CoupD(i,j)$, cuando $i \neq j$. Representa la relación entre el número de métodos de la clase j invocados por la clase i , al número total de métodos de la clase i , lo que indica el impacto de la clase j en la clase i (G. Gui & Scott, 2006). La expresión matemática se presenta a continuación:

$$CoupD(i,j) = \frac{|R_{i,j}|}{|R_i| + |M_i|} \quad (8)$$

Donde:

$R_{i,j}$ = Conjuntos de métodos y variables de instancia de la clase j invocados por la clase i

R_i = Relaciones de la clase

M_i = Métodos en la clase

Su rango de valores va de 0 y 1.

$CoupD(i, j)$, indica el acoplamiento directo entre la clase i y la clase j . Considerando que la clase i puede estar acoplada con más de una clase, se define una variante (9) de la métrica para medir el acoplamiento total de la clase i .

$$CoupT = \sum_{j=1}^{CBO} CoupD(i, j) \quad (9)$$

- **COF – CF Proporción de acoplamiento (Abreu y Melo, 1996)**

COF se define como la proporción entre el número máximo posible de acoplamientos en el sistema y el número real de acoplamientos no imputables a herencia. Indica la comunicación entre clases. Su rango de valores va de 0 y 1.

$$COF = \frac{\sum_{i=1}^{TC} [\sum_{j=1}^{TC} es_cliente(C_i C_j)]}{TC^2 - TC} \quad (10)$$

Donde:

TC = Total de clases

$$es_cliente = \begin{cases} 1 & \Leftrightarrow C_c \Rightarrow C_s \wedge C_c \neq C_s \\ 0 & \text{caso contrario} \end{cases}$$

La relación cliente-proveedor ($C_c \Rightarrow C_s$) representa que la clase cliente (C_c) contiene al menos una referencia no heredada de la clase proveedor (C_s).

De forma simplificada se representa de la siguiente forma (Robledo, 2004):

$$COF = \frac{\sum D}{TC^2 - TC} \quad (11)$$

Donde

D = Número de dependencias

TC = Total de clases

3.4 Análisis de regresión y de correlación

En Estadística Inferencial, se encuentra el análisis de regresión y correlación, utilizados en la Investigación Científica; cuando interesa analizar la relación entre dos o más variables,

medir la fuerza y expresar funcionalmente la relación mediante una función o modelo matemático (Orellana, 2008).

Matemáticamente hablando, una relación funcional está dada por (Mendiburu, 2007):

$$Y = f(x_1, \dots, x_n; \theta_1, \dots, \theta_m) \quad (12)$$

Donde:

Y = Variable respuesta (o dependiente)

x_i = La i -ésima variable independiente ($i = 1, \dots, n$)

θ_j = El j -ésimo parámetro en la función ($j = 1, \dots, m$)

f = La función

3.4.1. Tipos de regresión

Regresión Lineal Simple o Múltiple: Cuando la relación funcional entre la variable y (dependiente) y una (simple) o más (múltiple) variables x (independientes) sea una línea recta, representada por:

$$y = a + bx \quad (13)$$

Donde:

- El coeficiente a es la ordenada al origen, el punto donde la recta intercepta el eje y , es decir el valor de y cuando $x = 0$.
- El coeficiente b es la pendiente de la recta, mide el cambio en y por cada unidad de cambio en x . También conocido como coeficiente de regresión.

Regresión no lineal o curvilínea: Cuando la relación funcional entre variables no puede representarse mediante un modelo lineal, es preciso acudir a otras funciones que representen la relación de forma adecuada. Aquí podemos distinguir entre Regresión: Exponencial, Potencial, Logarítmica, Polinómica, etc.

Algunos problemas de regresión no lineal pueden linealizarse mediante una transformación en la formulación del modelo. Como en la regresión exponencial, potencial y logarítmica.

En la regresión polinomial, cuando la función (14) es de segundo grado, en ocasiones no se trata como un caso de regresión no lineal; la función es no lineal para x , pero lineal en función de los parámetros desconocidos a, b y c . Por lo tanto, los procedimientos computacionales para la regresión polinomial son procedimientos de regresión lineal (múltiple), con dos variables predictoras x y x^2 (Ramón, 2012).

$$y = ax^2 + bx + c \quad (14)$$

3.4.2. Método de mínimos cuadrados

Es una técnica de análisis numérico, en la que, dado un conjunto de pares ordenados (variable independiente, variable dependiente) y un conjunto de funciones, se intenta encontrar la función continua, que mejor se aproxime a los datos ("mejor ajuste"), de acuerdo con el criterio de mínimo error cuadrático.

En su forma más simple, intenta minimizar la suma de cuadrados de las diferencias en las ordenadas (llamadas residuos) entre los puntos generados por la función elegida y los correspondientes valores en los datos.

Ajustar una función de regresión significa buscar o definir la función que exprese con mayor precisión la relación entre las variables x y y .

- **Método de Mínimos Cuadrados para el Caso Polinomial**

La ecuación de un polinomio de grado n es (Salomón Ortiz Quintanilla, Jorge Segura Davila, 2005):

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{i=0}^n a_i x^i \quad (15)$$

Aplicándole el método de mínimos cuadrados. La curva propuesta es:

$$y_p = a_0 + a_1x + a_2x^2 + \dots + a_nx^n + e \quad (16)$$

Donde:

a_i = los coeficientes

e = es el error

Una estrategia es minimizar la suma de los cuadrados de los residuos S_r , entre la y medida y la y calculada con el modelo lineal, está dada por:

$$S_r = \sum e_i^2 = \sum (y_{i,medida} - y_{i,modelo})^2 = \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n)^2$$

Se realizan las derivadas parciales, dadas por:

$$\frac{\partial}{\partial a_0} S_r = \frac{\partial}{\partial a_0} \sum (y_i - a_0 - a_1 x_i - a_2 x_i^2 - \dots - a_n x_i^n)^2$$

$$\frac{\partial}{\partial a_0} S_r = -2 \sum (y_i - a_0 - a_1 x_i - a_2 x_i^2 - \dots - a_n x_i^n)$$

Y así sucesivamente hasta la n ecuación.

Se igualan a 0 las ecuaciones:

$$\frac{\partial}{\partial a_0} S_r = 0$$

Reordenando, para desarrollar el siguiente sistema de ecuaciones normales:

$$\begin{aligned} a_0(m) + a_1 \sum x_i + a_2 \sum x_i^2 + \dots + a_n \sum x_i^n &= \sum y_i \\ a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 + \dots + a_n \sum x_i^{n+1} &= \sum y_i x_i \\ a_0 \sum x_i^n + a_1 \sum x_i^{n+1} + a_2 \sum x_i^{n+2} + \dots + a_n \sum x_i^{n+2} &= \sum y_i x_i^2 \end{aligned}$$

Todas las sumatorias son desde $i = 1$ hasta m (donde m es el número de puntos).

Podemos escribir el sistema de ecuaciones normales obtenido en la forma:

$$S_x a = S_{xy} \tag{17}$$

$$\begin{bmatrix} m & \sum x & \sum x^2 & \dots & \sum x^n \\ \sum x & \sum x^2 & \sum x^3 & \dots & \sum x^{n+1} \\ \sum x^n & \sum x^{n+1} & \sum x^{n+2} & \dots & \sum x^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_n \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum yx \\ \sum yx^n \end{bmatrix} \tag{18}$$

Donde:

S_x = Matriz de sumatorias de potencias de x

a = Vector de coeficientes. Las constantes del polinomio

S_{xy} = Vector de sumatorias de potencias de x con y 's

Para obtener el error estándar estimado es:

$$S_{y/x} = \sqrt{\frac{S_r}{m - (n + 1)}} \quad (19)$$

3.4.3. Coeficiente de correlación

La correlación indica la fuerza y la dirección de una relación y proporcionalidad entre dos variables estadísticas. Se considera que dos variables cuantitativas están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores homónimos de la otra: si tenemos dos variables (a y b) existe correlación si al aumentar los valores de a lo hacen también los de b y viceversa (Ramón, 2012). La correlación entre dos variables no implica, que existe necesariamente una relación causa-efecto entre ellas (Alejandro Rosas Snell, 2013).

El objetivo del análisis de correlación de Pearson es medir la intensidad de la asociación lineal observada entre dos variables cuantitativas. Su valor varía entre -1 y $+1$; esto es (Mendiburu, 2007):

$$-1 \leq r \leq 1$$

Si $r = -1$, la asociación es perfecta pero inversa o negativa; es decir, cuando al aumentar la medida de una de las variables, los valores de la otra disminuyen.

Si $r = +1$, la asociación también es perfecta pero directa o positiva, es decir cuando al aumentar la medida de una de las variables, la otra variables también aumenta.

Si $r = 0$, no existe asociación entre las dos variables, correlación nula.

El coeficiente de correlación se puede clasificar de la siguiente manera, según su valor o rango:

- Perfecta cuando $r = 1$
- Excelente cuando $r = 0.9 \leq r < 1$
- Buena cuando $r = 0.8 \leq r < 0.9$
- Regular cuando $r = 0.5 \leq r < 0.8$
- Mala cuando $r < 0.5$

El coeficiente de correlación de Pearson está dado por:

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}} \quad (20)$$

3.4.4. Coeficiente de determinación

Mide el porcentaje de variación en la variable respuesta (dependiente), explicada por la variable independiente (Mendiburu, 2007).

El coeficiente de determinación representado por R^2 o r^2 , nos dice qué porcentaje de la variabilidad total en la variable y (dependiente) puede ser explicada por la variable x (independiente), en consecuencia, es una medida de la capacidad de *predicción* del modelo elegido (Orellana, 2008).

Su valor varía entre $0 \leq R^2 \leq 1$. Por tanto (Mendiburu, 2007):

Si $R^2 = 1$, entonces no hay residuos, habrá una dependencia funcional perfecta. Cuanto más se acerque dicho valor a la unidad, mayor poder explicativo tendrá el modelo de regresión.

Si $R^2 = 0$, la variable x no explica en absoluto ninguna de las variaciones de la variable y , de modo que, o bien el modelo es inadecuado, o bien las variables son independientes. Cuanto más cercano a 0 esté dicho valor, menor poder explicativo.

El Coeficiente de determinación está dado por:

$$R^2 = \frac{SCR}{SCE} \quad (21)$$

Donde:

SCR = Suma de cuadrados debido a la regresión $\sum(\hat{y}_i - \bar{y})^2$

SCE = Suma de cuadrados debido al error $\sum(y_i - \hat{y}_i)^2$

Capítulo 4: Diseño Experimental

4.1 Formulación de preguntas e hipótesis

4.1.1. Pregunta de investigación

¿Cuál es el sentido y grado de influencia que tienen entre sí las métricas de calidad del software, cohesión y acoplamiento en diferentes unidades de programa que participan colaborativamente para satisfacer una meta de valor u objetivo?

4.1.2. Formulación de hipótesis

“Una hipótesis es una declaración o explicación que se sugiere por el conocimiento o la observación que, sin embargo, no ha sido probada o desmentida” (Abreu, 2012).

Las hipótesis de investigación o hipótesis de trabajo, se definen como *“proposiciones tentativas acerca de las posibles relaciones entre dos o más variables”* (Hernández Roberto, 2004).

En el contexto de este trabajo de tesis se plantean las hipótesis de la siguiente manera:

Hipótesis general

- En unidades de programa que participan colaborativamente para cumplir una meta de valor, el nivel de cohesión influye en el nivel de acoplamiento

Hipótesis específicas

- Unidades de programa con alto nivel de cohesión producen bajo acoplamiento
- Unidades de programa con bajo nivel de cohesión producen alto acoplamiento
- Unidades de programa con alto nivel de cohesión producen alto nivel de acoplamiento
- Unidades de programa con bajo nivel de cohesión producen bajo nivel de acoplamiento

Hipótesis estadística

Para aceptar o rechazar la hipótesis general y las hipótesis específicas, mediante el análisis estadístico de los datos, se definen las hipótesis en términos estadísticos:

- *Hipótesis nula* (H_0): El grado de influencia o relación de las métricas de cohesión y de acoplamiento es nulo en unidades de programa.
- *Hipótesis alternativa* (H_1): El grado de relación de la medida de acoplamiento es inversamente proporcional a la medida de cohesión.
- *Hipótesis alternativa* (H_2): El grado de relación de la medida de acoplamiento es directamente proporcional a la medida de cohesión.

4.2 Identificación de variables

Para efectos de este estudio experimental, se han identificado 2 tipos de variables: las variables dependientes y las variables independientes.

4.2.1 Variables independientes

Son las variables que pueden cambiar libremente su valor y que no dependen de otra variable, se identifica como la causa o antecedente. Esta variable es manipulada por el investigador. Para este estudio experimental se identificó como variable independiente, la medida de cohesión en unidades de programa conocidos como objetos de prueba.

- “*Medida de cohesión*”: Para obtener esta medida se utilizaron algunas de las métricas de la literatura, como LCOM*, LCOM5, ClassCoh, WTCoh y Coh. El cálculo de estas métricas se realiza de forma automática, mediante el marco orientado a objetos para la medición de la calidad de arquitecturas de software, desarrollado en esta tesis.

4.2.2 Variables dependientes

Se considera como el efecto o resultado que se observa en el estudio experimental, estas variables son influenciadas por los valores de las variables independientes (de entrada). Para este estudio experimental se identificó como variable dependiente el acoplamiento:

- “*Medida de acoplamiento*”: Medida que se obtiene mediante las métricas CoupT, WTCoup y COF. El cálculo de estas métricas se hace de forma automática mediante

el marco orientado a objetos para la medición de la calidad de arquitecturas de software, desarrollado en esta tesis. Esta medida cambia con respecto a la medida de cohesión.

4.3 Diseño del plan de pruebas

“El plan de pruebas describe el alcance y enfoque de las actividades de pruebas previstas, e identifica las características a ser probadas” (Elaine, 2000). Las pruebas de este experimento tienen como alcance obtener la medición de las métricas de cohesión y acoplamiento a nivel de clase y a nivel de sistema, en cada uno de los casos de pruebas.

A continuación, se describe el plan de pruebas para el estudio experimental:

- Casos de prueba: Son unidades experimentales sobre las cuales se aplicará el experimento con el fin de comprobar el comportamiento que tienen las variables de entrada (Cohesión) sobre la variable de salida (Acoplamiento). Como se menciona en las limitaciones de esta investigación, únicamente se tomaron como casos de prueba los referidos en el trabajo de antecedente (Rogel, 2014).
- Diseño de los casos de prueba: Con el fin de manipular la variable independiente (medida de cohesión) y conocer la influencia que tiene sobre la variable dependiente (medida de acoplamiento); cada uno de los casos de prueba se diseñó con una estructura o arquitectura interna diferente. En el diseño se realizaron variaciones en su arquitectura en las cuales se pudiera dar el efecto requerido. Un punto que se consideró, es que el cambio en la arquitectura no debe afectar el funcionamiento del caso de prueba y que la manipulación no se basa en mejorar el diseño de la arquitectura o lo contrario.
- Se establecieron las métricas de cohesión y acoplamiento que se aplicaron a los casos de prueba. Ver capítulo 3.
- Se desarrolló una herramienta computacional para el cálculo automático de las métricas de cohesión y acoplamiento.

4.3.1 Descripción de los casos de prueba

Para este experimento se requiere programas en código Java. A continuación, se describen brevemente los programas que se tomaron como casos de prueba. Para profundizar más, consultar el trabajo de tesis (Rogel, 2014).

Caso de Prueba 1: Generación de series de Bernoulli

Su objetivo es calcular un conjunto de números de Bernoulli. Los números de Bernoulli (B_n) constituyen una sucesión de números racionales con conexiones profundas en la teoría de números. La serie de Bernoulli se genera de forma inductiva, donde el primer número de Bernoulli (por definición 1) genera al segundo número de la serie, el tercer número de Bernoulli se genera a partir del primero y del segundo y así sucesivamente.

Caso de Prueba 2: Generación de variables aleatorias continuas

Su objetivo es generar un conjunto de variables aleatorias continuas mediante las distribuciones de densidad de Gauss, Gamma y Beta.

La gráfica de la distribución de densidad normal tiene una forma acampanada y es simétrica respecto de un determinado parámetro estadístico. Esta curva se conoce como campana de Gauss y es el gráfico de la función gaussiana. La distribución de densidad Gamma es una distribución de probabilidad continua con dos parámetros k y λ . La distribución de densidad Beta es una distribución de probabilidad continua con dos parámetros α y b .

Caso de Prueba 3: Generación de variables aleatorias discretas

Su objetivo es generar un conjunto de variables aleatorias discretas mediante algunas distribuciones de probabilidad. El caso de prueba calcula un conjunto de variables aleatorias discretas utilizando distribuciones de probabilidad: Binomial, Binomial Negativo, Poisson, Geométrica, HiperGeométrica, Multinomial, y Uniforme Discreta.

Caso de prueba 4: Multiplicación de matrices mediante el algoritmo de Hipercubo simple

Su objetivo es realizar la multiplicación de 2 matrices de dimensión $n \times n$ utilizando el algoritmo de Hipercubo.

El algoritmo de multiplicación de matrices consta de 3 pasos: el primer paso es distribuir los valores de las matrices A y B en los nodos del hipercubo, el segundo paso realiza el producto de los valores entre los vértices de los cubos y el tercer paso realiza la suma de los valores de los vértices, obteniendo así el resultado final.

Caso de Prueba 5: Ordenamiento de series de Taylor mediante Bucket-Sort

Su objetivo es generar un conjunto de elementos mediante una serie de Taylor para posteriormente ordenarlos utilizando el ordenamiento por casilleros.

Este caso de prueba consiste de 3 pasos: 1) la creación de una serie de Taylor, 2) la ejecución del algoritmo de ordenamiento por casilleros (Bucket-Sort) y 3) un algoritmo de ordenamiento interno. El ordenamiento por casilleros distribuye los elementos en contenedores, donde cada contenedor puede almacenar únicamente un conjunto de elementos

que cumpla con cierta condición. Una vez almacenados cada uno de los elementos en su respectivo casillero, se recurre a algún algoritmo de ordenamiento para ordenarlos internamente.

4.3.2 Diseño de los casos de prueba

Cada caso de prueba fue estructurado al menos con 12 arquitecturas diferentes. Las arquitecturas son definidas en función de un diseño orientado a objetos. Cada arquitectura tiene un diseño y estructura interna orientada a objetos única, la cual fue re-factorizada manualmente en arquitecturas alternativas con varios niveles de cohesión, para observar la variación del nivel de acoplamiento en cada una de estas arquitecturas.

Un requisito que se mencionó anteriormente es que las diferentes arquitecturas obtenidas en el proceso de refactorización sigan conservando el comportamiento funcional de la arquitectura original.

El diseño de la arquitectura de algunos casos de prueba inicia con una arquitectura con un nivel de granulación medio, es decir una clase con varios métodos, variando la arquitectura hasta obtener una arquitectura de granulación fina, es decir un conjunto de clases con uno o varios métodos. Desde la primera arquitectura hasta la final, se realizaron varios cambios en las arquitecturas, desde el aumento o disminución de métodos, atributos e instancias de objetos. Para efectos de simplificar el desarrollo de este apartado, sólo se describen las variaciones de la arquitectura de uno de los casos de prueba.

Caso de prueba: Hiper cubo

Escenario#1: Esta arquitectura está compuesta por una clase con varios métodos (Figura 1), su nivel de granulación es medio. En esta arquitectura, se puede observar que la medida de acoplamiento es cero, la clase no requiere de otras clases para cumplir con su meta de valor. Lo que pudiera significar que la clase es suficiente y completa para cumplir con una única responsabilidad.

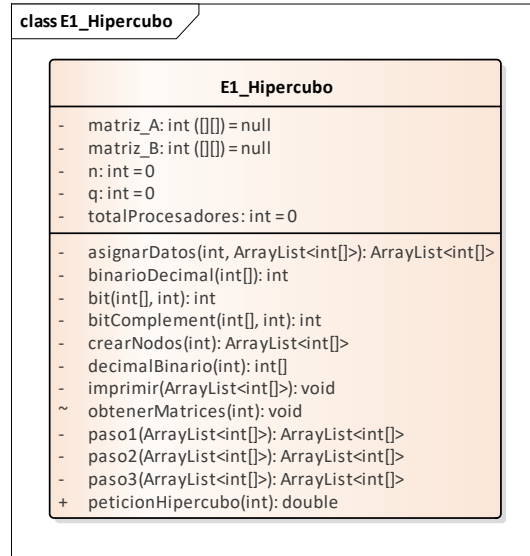


Figura 1. Arquitectura del escenario#1

Escenario#2: Se modificó la arquitectura del escenario #1 (Figura 1). Para esto se creó una nueva clase con el método imprimir, provocando un canal de comunicación (Figura 2). La clase E2_Hipercubo depende de otra clase para cumplir con su meta de valor. Lo que pudiera significar que la clase es insuficiente e incompleta para cumplir con una única responsabilidad.

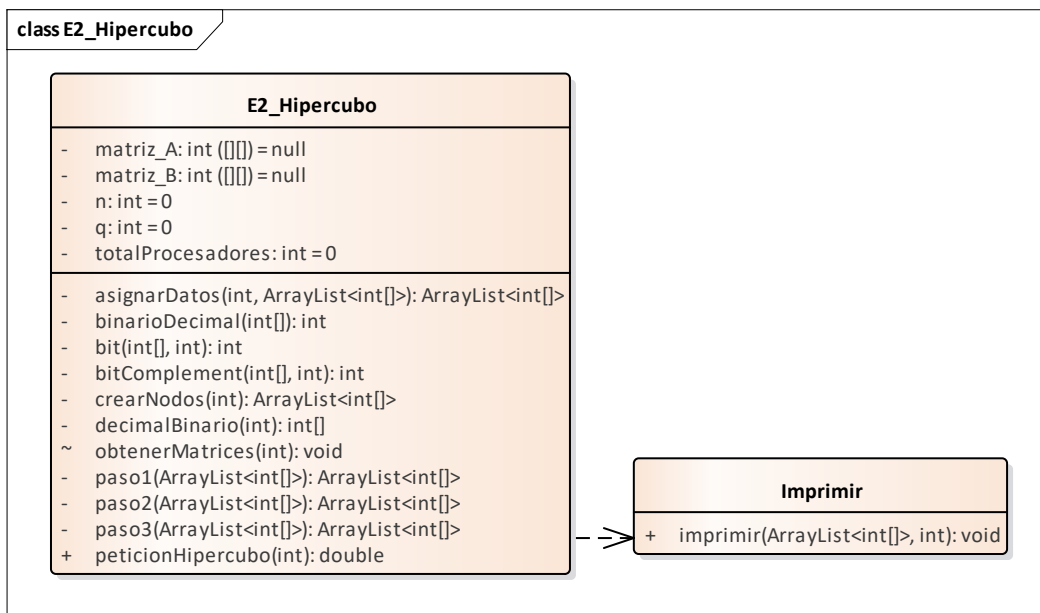


Figura 2. Arquitectura del escenario#2

Escenario#3: Se modificó la arquitectura del escenario #2 (Figura 2). Para esto se creó una nueva clase con el método “obtener matriz”, provocando el incremento de los canales de comunicación (Figura 3). La clase E3_Hipercubo depende de otra clase para cumplir con su meta de valor.

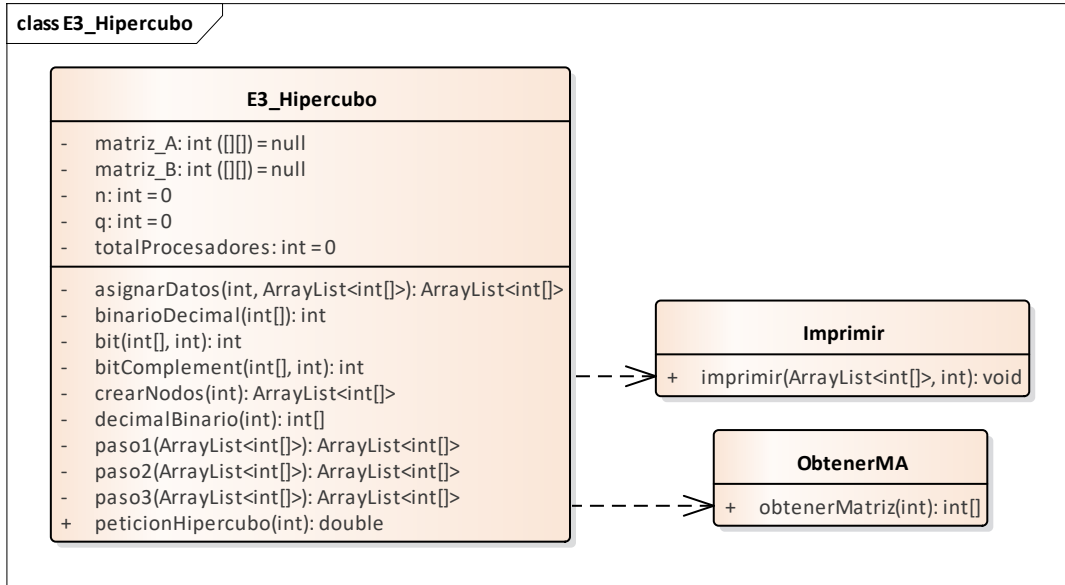


Figura 3. Arquitectura del escenario#3

Escenario#4: Se modificó la arquitectura del escenario #3 (Figura 3). Para esto se creó una nueva clase con el método “crearNodo”, provocando el incremento de los canales de comunicación. La dependencia de la clase E4_Hipercubo cada vez es mayor (Figura 4).

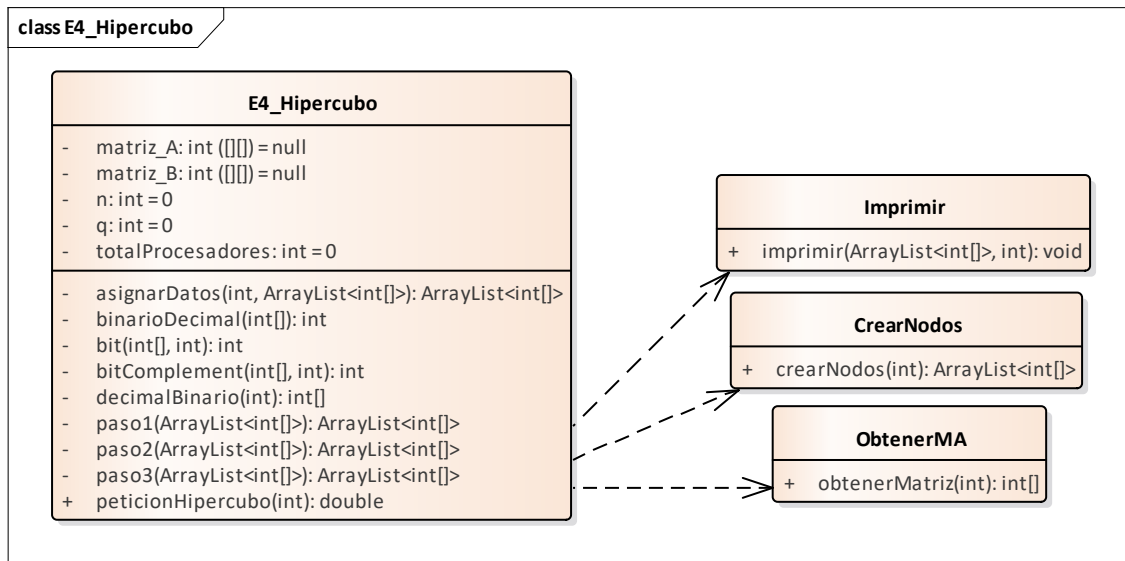


Figura 4. Arquitectura del escenario#4

Escenario#5: Se modificó la arquitectura del escenario #4 (Figura 4). Para esto se creó una nueva clase “DecimalBinario”. La instancia del objeto se hizo a nivel de clase (relación de asociación), incrementando el número de atributos (Figura 5). Las diferentes arquitecturas muestran una disminución en el número de métodos en la clase E5_Hipercubo, pero un aumento en los canales de comunicación.

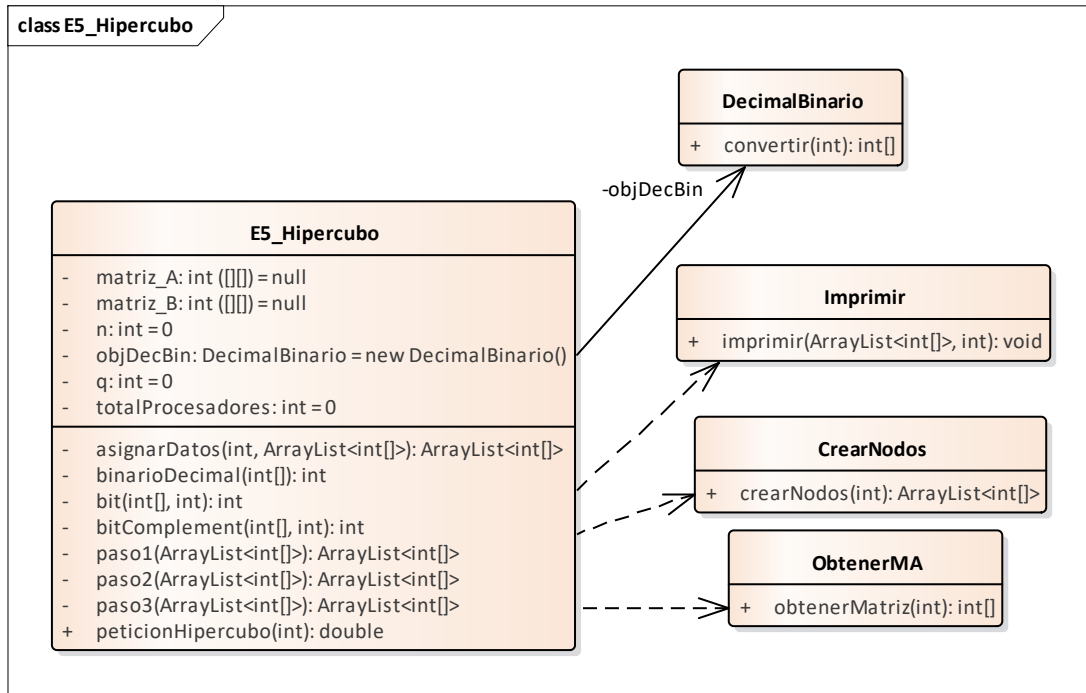


Figura 5. Arquitectura del escenario#5

Escenario#6: Se modificó la arquitectura del escenario #5 (Figura 5). Para esto el método “bitComplement” se colocó en la clase “DecimalBinario” (Figura 6). Esta arquitectura muestra una disminución en los métodos de la clase “E6_Hipercubo” y un aumento en la clase “DecimalBinario”. Existe una dependencia en ambos sentidos entre estas clases.

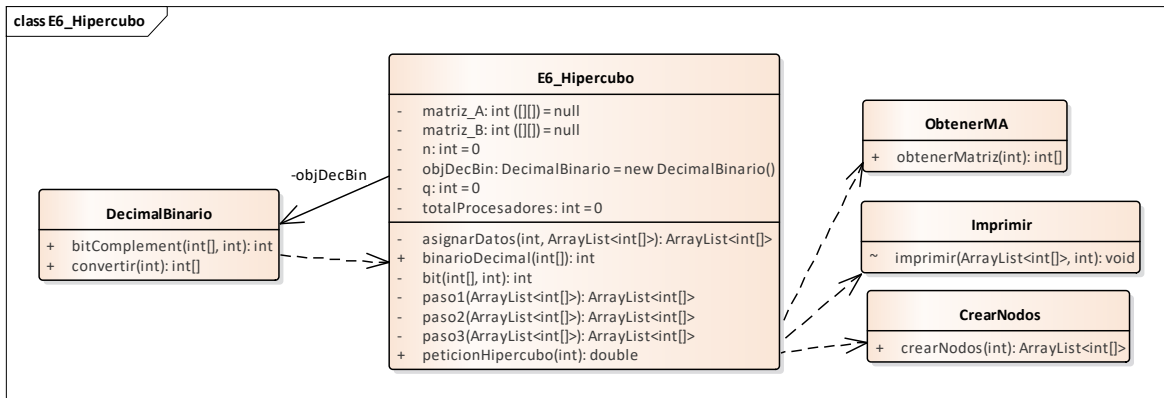


Figura 6. Arquitectura del escenario#6

Escenario#7: Se modificó la arquitectura del escenario #6 (Figura 6). El método “binarioDecimal” se colocó en la clase “DecimalBinario” (Figura 7). Con este cambio se reduce el número de dependencia, entre esta clase y la clase “E7_Hipercubo”, dejó de ser en ambos sentidos.

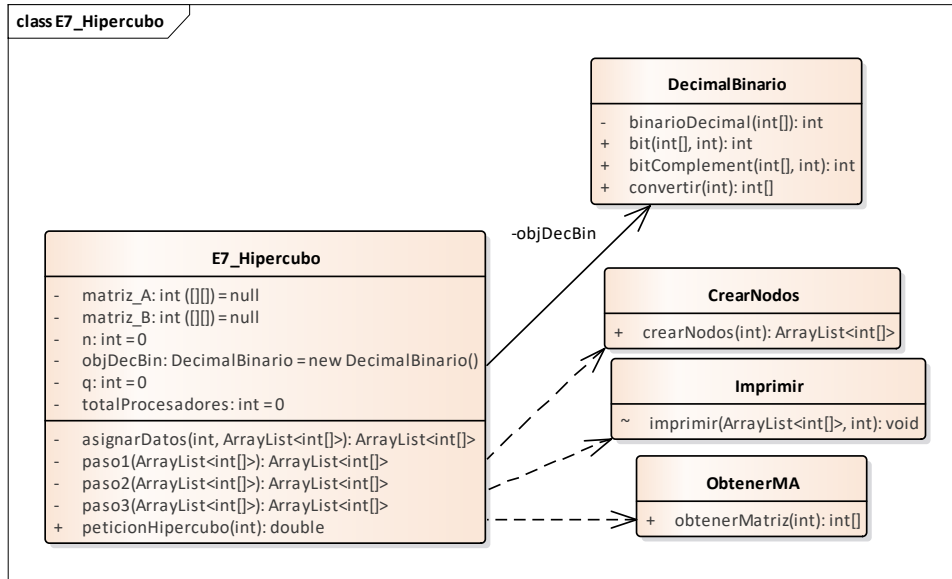


Figura 7. Arquitectura del escenario#7

Escenario#8: Se modificó la arquitectura del escenario #7 (Figura 7). Para esto se crearon las clases “AsignarDatos” y “Algoritmo”. La clase “Algoritmo” contiene los datos de inicialización (Figura 8). Este cambio aumenta los canales de comunicación y disminuyen los atributos y métodos de la clase “E8_Hipercubo”.

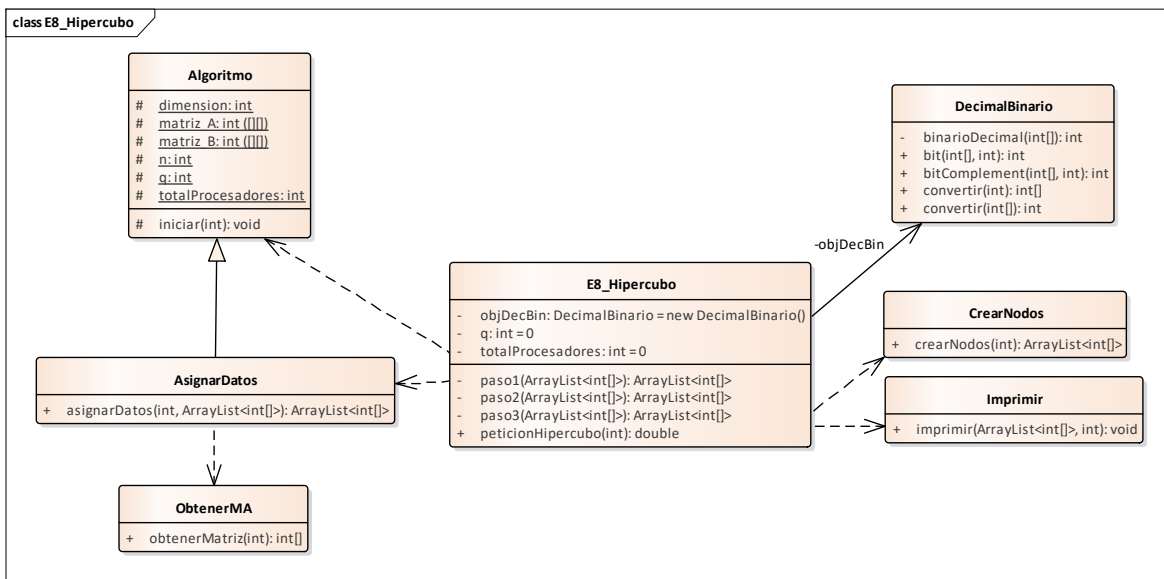


Figura 8. Arquitectura del escenario#8

Escenario#9: Se modificó la arquitectura del escenario #8 (Figura 8). Para esto se crearon las clases “Nodo” y “Paso3”. Este cambio aumenta los canales de comunicación y disminuyeron los métodos en la clase principal (Figura 9).

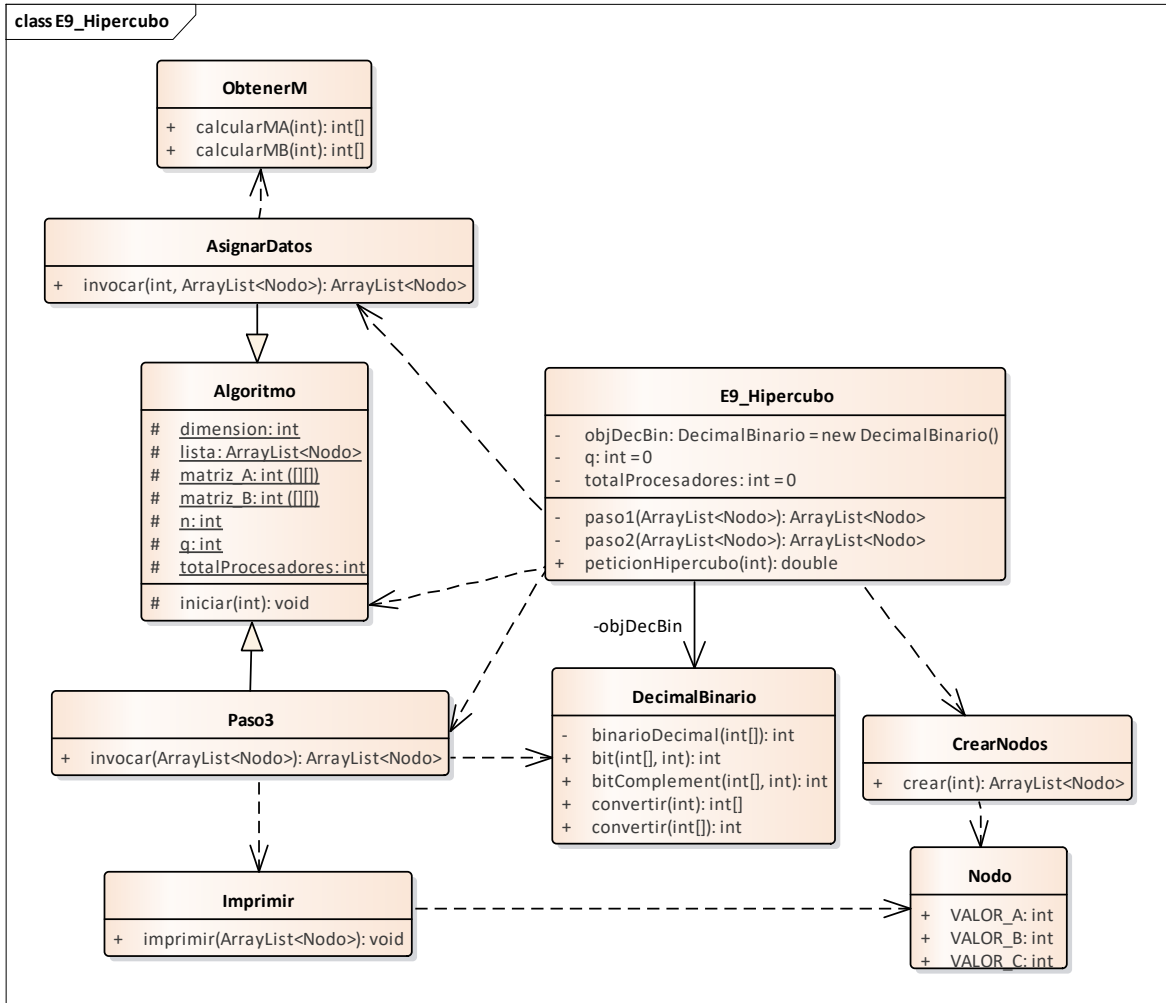


Figura 9. Arquitectura del escenario#9

Escenario#10: Se modificó la arquitectura del escenario #9 (Figura 9). Para esto se creó una nueva clase con nombre “Paso2” (Figura 10). La variación de esta arquitectura es similar a la anterior. Incrementó el número de canales y el número de clases en el paquete.

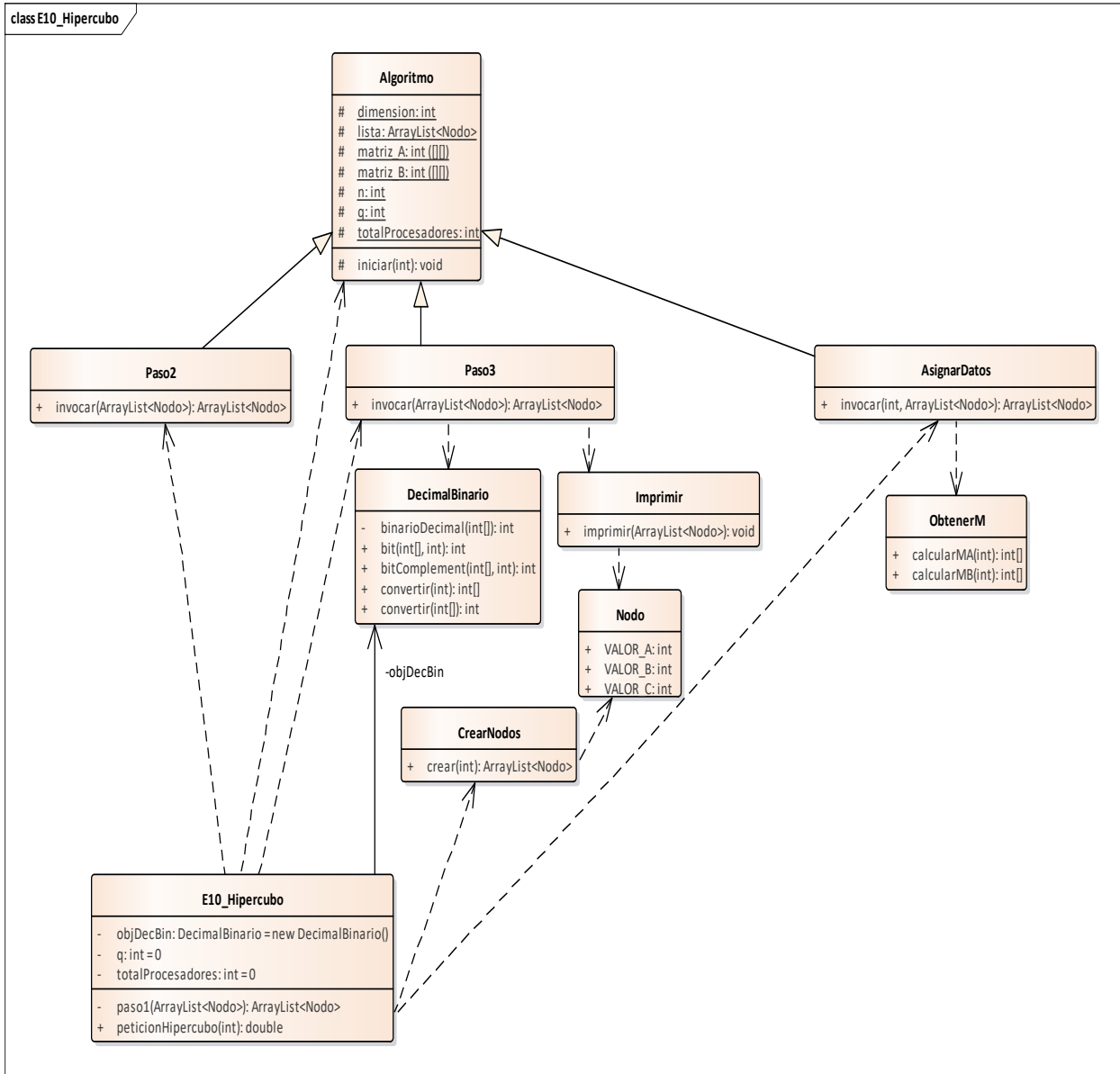


Figura 10. Arquitectura del escenario#10

Escenario#11: Se modificó la arquitectura del escenario #10 (Figura 10). Se creó una nueva clase “Paso1”. Disminuyó el número de atributo en la clase “E11_Hipercubo”. El método principal de esta clase se dividió en dos métodos (Figura 11). En este caso se incrementó el número de canales de comunicación y el número de clases en el paquete.

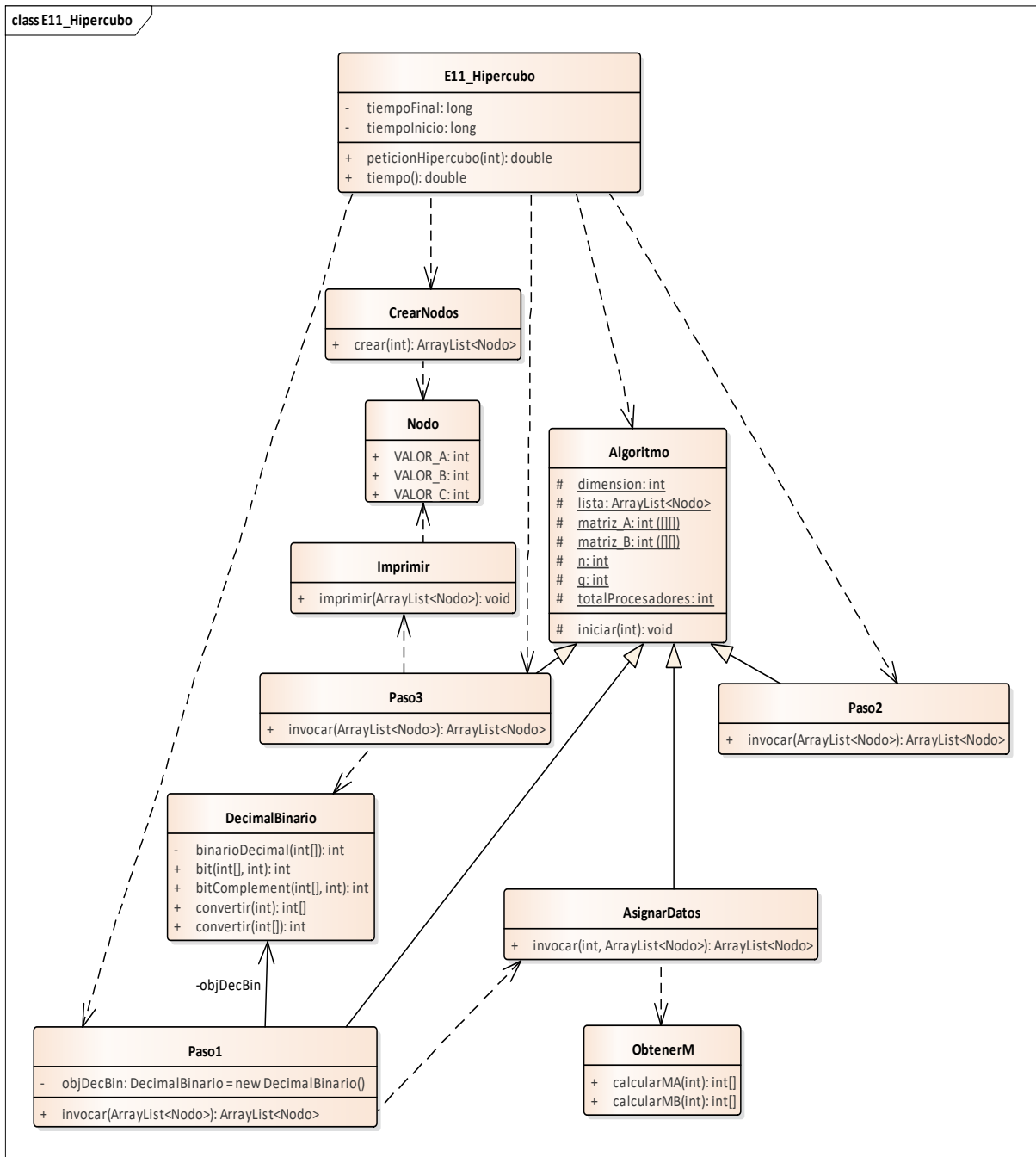


Figura 11. Arquitectura del escenario#11

Escenario#12: Se modificó la arquitectura del escenario #11 (Figura 11). La invocación de métodos de la clase “E12_Hipercubo” se pasaron a la clase “Algoritmo” (Figura 12). Esta arquitectura tiene un nivel de granulación fino, un conjunto de clases con pocos métodos.

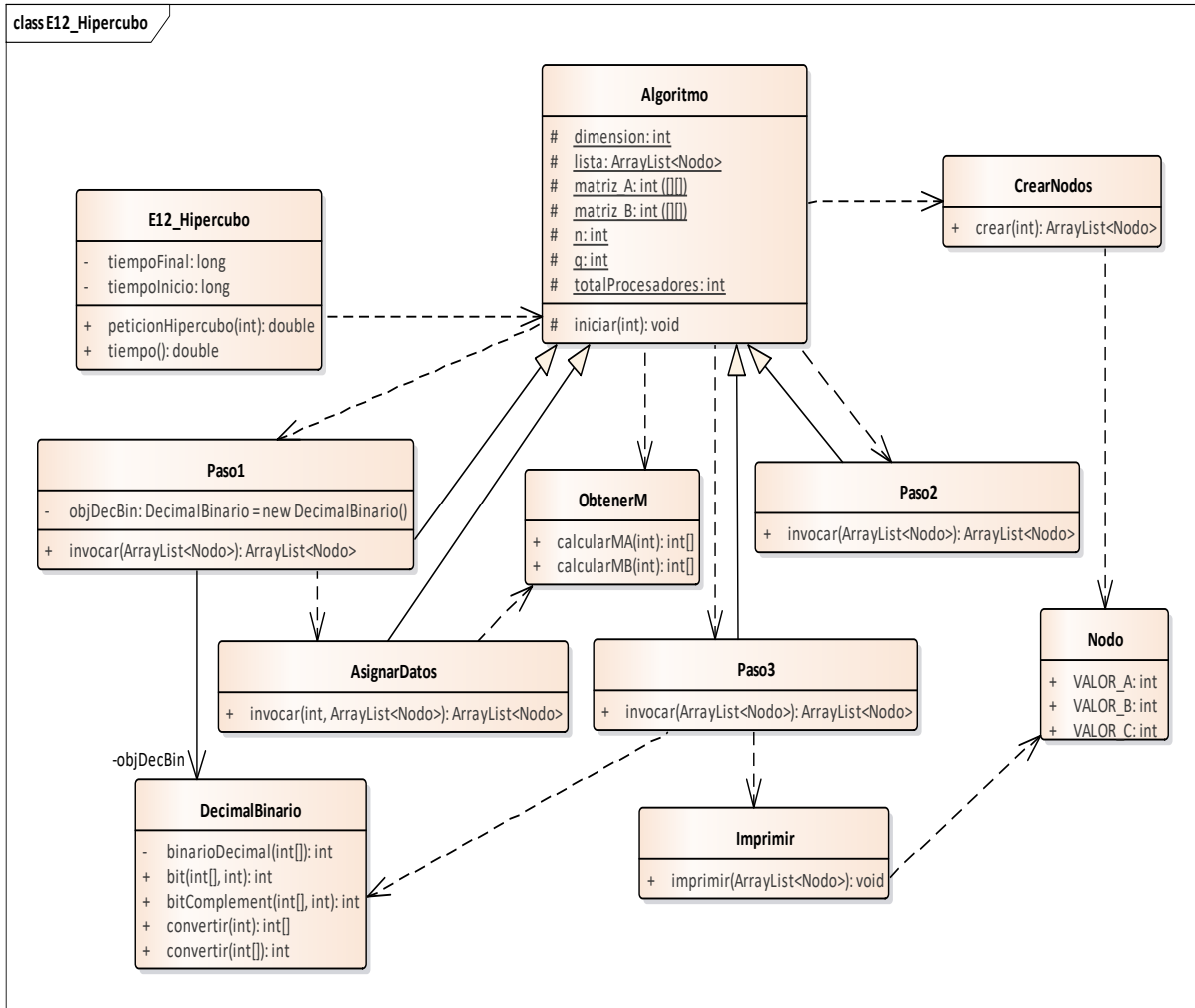


Figura 12. Arquitectura del escenario#12

Siguiendo el patrón de variación representado en los escenarios anteriores, se incrementa el número de clases y en consecuencia el número de canales de comunicación, siendo en el mejor de los casos, n-1 canales de comunicación, se podría continuar con el diseño de más escenarios hasta lograr una arquitectura con niveles de granulación muy fina, es decir un conjunto de clases, cada una con un único método. En algunos de los casos de prueba si se llegó a este nivel de granulación.

4.3.3 Desarrollo del instrumento de pruebas

Para facilitar el cálculo de las métricas de cohesión (LCOM*, LCOM4, LCOM5, ClassCoh, Coh, WTCoh) y acoplamiento (CBO, COF, WTCoup), se implementó un marco orientado a objeto para la medición de calidad en arquitecturas de software. Este marco orientado a objeto fue desarrollado utilizando el lenguaje java y el metalenguaje ANTLR (ANother Tool for Language Recognition).

ANTLR proporciona la capacidad para construir reconocedores (parsers), intérpretes, compiladores y traductores de lenguajes, a partir de las descripciones gramaticales de los mismos (las cuales contienen acciones semánticas), facilitando la obtención de los datos necesarios para el cálculo de las métricas.

La arquitectura del marco orientado a objetos está organizada en subárboles. En un subárbol están las métricas de acoplamiento y en otro subárbol están las métricas de cohesión. Cada subárbol incluye el cálculo de las métricas con algoritmos estratégicos con comportamientos diferentes. La arquitectura está diseñada conforme a los principios de diseño orientado a objetos: abierto/cerrado, inversión de dependencias, sustitución, única responsabilidad y no-repetición. Esta arquitectura garantiza el desarrollo del marco orientado a objetos con las mejores prácticas de la ingeniería de software para efectos de su mantenimiento y control de su evolución en el tiempo.

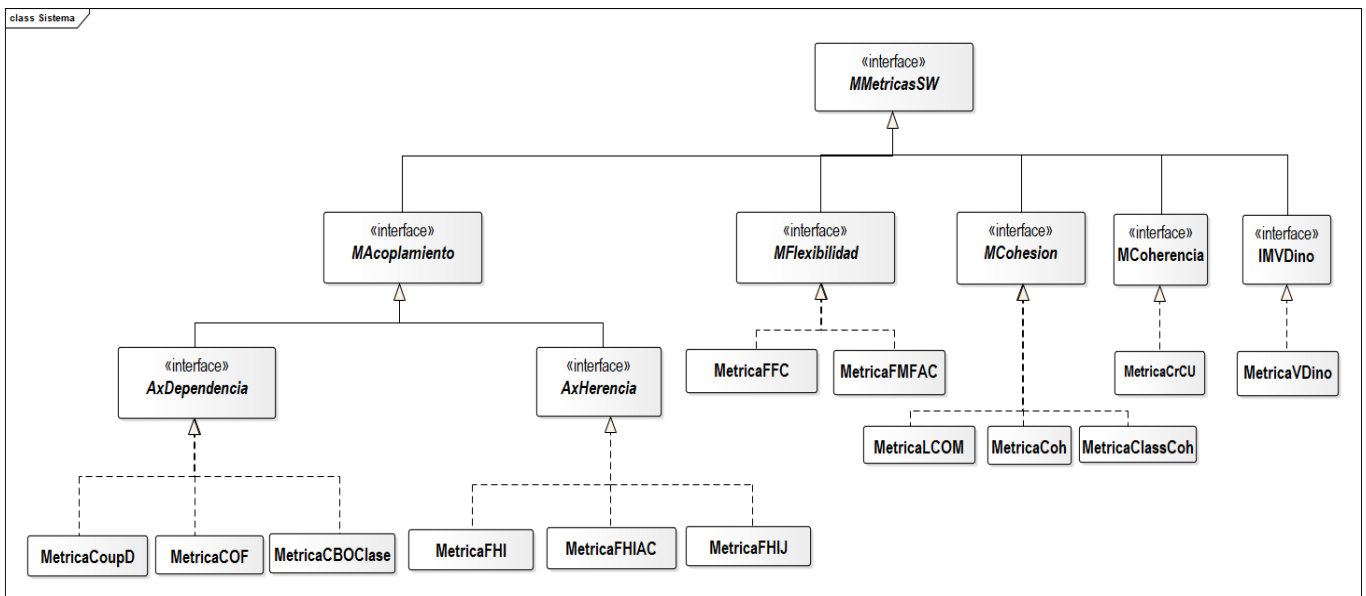


Figura 13. Arquitectura de clase del Marco de métricas orientado a objetos

Capítulo 5: Ejecución del Experimento y Recolección de Mediciones

5.1 Ejecución del experimento

Para identificar los 5 casos de prueba, en el resto del documento se hace referencia a cada uno de ellos como Bernoulli, Continuas, Discretas, Hipercubo y TaylorBucket.

La ejecución de las métricas de cohesión y acoplamiento para cada caso de prueba se clasificaron de la siguiente manera:

- **Mediciones a nivel de clase:** En esta sección se muestran las mediciones de las métricas de cohesión (LCOM*, LCOM4, LCOM5, ClassCoh, Coh) y acoplamiento (CBO, CoupT) para cada caso de prueba.
- **Mediciones a nivel de sistema o paquete:** En esta sección se muestran las mediciones de las métricas de cohesión (LCOM*, LCOM5, WTCoh, Coh) y acoplamiento (COF, WTCoup) para cada caso de prueba.

Cabe mencionar que las mediciones se obtuvieron con la herramienta desarrollada (Marco de métricas orientado a objetos).

Las métricas de cohesión solo se pueden medir cuando la clase tiene uno o más atributos y más de un método. Existen valores predefinidos para LCOM5 y Coh, cuando no cumplen con esta restricción (Al Dallal, 2011). La tabla 2 muestra los valores predefinidos, donde “m” es el número de métodos, “a” el número de atributos y “AxM” es la suma del número de atributos distintos a los que accede cada método de la clase.

Tabla 2. Valores predefinidos para los casos que no cumplen con la restricción.

Métrica	m=0 y a>0	m=1 y a=0	m=1 y a>0	m>1 y a=0
LCOM5	0	0	1	1
Coh	1	1	AxM/a	0

Para las métricas LCOM* y ClassCoh, no se establecen valores predefinidos para los casos que no cumplen con la restricción de medición. Por lo tanto, en las tablas 3-12 se muestran mediciones invalidas representadas con el signo “?”.

5.2 Recolección de mediciones

En las tablas 3-12 se muestran las mediciones de las métricas de cohesión y acoplamiento, correspondientes para cada caso de prueba, tanto a nivel de clase como a nivel de sistema. Las mediciones marcadas en color verde representan el mejor valor y el color rojo el peor valor de cada métrica.

- **Mediciones a nivel de clase:** Las mediciones se realizan sobre la clase principal de cada caso de prueba.

Tabla 3. Mediciones a nivel de clase CP. Bernoulli

CP. Bernoulli																	
#Esce	Clase	#Met	#Atrib	ΣAxM	LCOM*		LCOM4		LCOM5		Coh		ClassCoh		CBO	CoupT	
					0 bueno 1 malo	1 bueno >1 malo	0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo	1 bueno 0 malo	0 bajo 1 alto						
1	E1_Bernoulli	1	0	0	?	1	0.00	1.00	?	0	0.00						
2	E2_Bernoulli	1	2	2	?	1	1.00	1.00	?	0	0.00						
3	E3_Bernoulli	2	0	0	?	1	1.00	0.00	0.000	0	0.00						
4	E4_Bernoulli	2	2	3	0.50	1	0.33	0.75	0.500	0	0.00						
5	E5_Bernoulli	11	2	3	0.95	1	0.90	0.14	0.009	0	0.00						
6	E6_Bernoulli	2	2	2	1.00	1	0.67	0.50	0.000	2	0.88						
7	E7_Bernoulli	3	2	4	0.50	1	0.40	0.67	0.333	2	0.83						
8	E8_Bernoulli	4	2	3	0.83	1	0.71	0.38	0.083	1	0.78						
9	E9_Bernoulli	5	2	5	0.62	1	0.56	0.50	0.250	1	0.74						
10	E10_Bernoulli	3	3	5	0.67	1	0.50	0.56	0.222	2	0.83						
11	E11_Bernoulli	2	2	3	0.50	1	0.33	0.75	0.500	7	0.91						
12	E12_Bernoulli	3	2	3	0.75	1	0.60	0.50	0.167	6	0.86						

En la tabla 3, se muestran las mediciones de las métricas de cohesión y acoplamiento para el caso de prueba Bernoulli. En los primeros escenarios (#Esce.1-3) se muestran los casos donde las métricas LCOM* y ClassCoh no pueden ser calculadas (“?”), mismos casos donde las métricas LCOM5 y Coh tienen valores predefinidos.

Para este caso de prueba el mejor valor de las métricas LCOM* y ClassCoh se ve en los escenarios (#Esce.4 y 11), este valor se obtiene cuando al menos dos métodos comparten atributos (ΣAxM), es decir entre más métodos compartan atributos la medida de cohesión será mejor. El peor valor se obtiene cuando los métodos no comparten ningún atributo.

En todos los escenarios de la tabla 3, la métrica LCOM4 presenta un valor bueno (1), esto significa que todos los elementos de la clase (métodos y atributos), están colaborando para atender una única responsabilidad. Las métricas LCOM5 y Coh presentan su mejor valor,

cuando la clase tiene un solo método y cero atributos (#Esce. 1). El peor valor cuando la clase tiene más de un método y cero atributos (#Esce. 3), significa que los métodos de la clase no comparten atributos, por lo que pueden ser separados en clases diferentes, mejorando el valor de cohesión. Sin embargo, estos métodos podrían estar relacionados mediante la invocación. Este tipo de relación no se considera en ninguna de las métricas de cohesión. Adicionalmente, existe una contradicción en los valores predefinidos en las métricas LCOM5 y Coh, cuando la clase tiene solo un método y uno o más atributos (#Esce. 2); para LCOM5 representa un valor malo; si la clase tiene solo un método, no hay razón para tener atributos en la clase, y para Coh representa un valor bueno; considerando que todos los atributos son usados por uno o más métodos. Y finalmente en la métrica CoupT se presenta un decremento en su valor cuando el número de métodos en la clase aumenta.

Tabla 4. Mediciones CP. V.A. Continuas

CP. Continuas												
#Esce	Clase	#Met	#Atrib	ΣAxM	LCOM*		LCOM5	Coh		ClassCoh	CBO	CoupT
					0 bueno 1 malo	1 bueno >1 malo	0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo	0 bajo 1 alto		
1	E1_Continuas	1	0	0	?	1	0.00	1.00	?	4	0.95	
2	E2_Continuas	2	1	2	0.00	1	0.00	1.00	1.000	4	0.91	
3	E3_Continuas	3	1	2	0.50	1	0.50	0.67	0.333	3	0.80	
4	E4_Continuas	3	1	2	0.50	1	0.50	0.67	0.333	5	0.87	
5	E5_Continuas	4	1	2	0.67	1	0.67	0.50	0.167	4	0.82	
6	E6_Continuas	4	2	3	0.83	1	0.71	0.38	0.083	4	0.81	
7	E7_Continuas	3	4	5	0.88	1	0.64	0.42	0.167	5	0.86	
8	E8_Continuas	4	3	4	0.89	1	0.73	0.33	0.167	4	0.75	
9	E9_Continuas	5	5	6	0.95	1	0.79	0.24	0.100	3	0.67	
10	E10_Continuas	6	7	8	0.97	1	0.83	0.19	0.067	3	0.62	
11	E11_Continuas	7	7	8	0.98	1	0.85	0.16	0.048	1	0.46	
12	E12_Continuas	6	1	1	1.00	1	1.00	0.17	0.000	1	0.50	

En la tabla 4, se muestran las mediciones de las métricas de cohesión y acoplamiento para el caso de prueba V.A. Continuas. En los primeros escenarios (1-5) se mantuvo fijo el número de atributos y se varió el número de métodos en la clase, provocando variación en la medida de cohesión y acoplamiento. En los escenarios (3 y 4), solo se varió el número de canales de comunicación de la clase (CBO); la medida de acoplamiento cambió, pero la medida de cohesión se mantuvo. Por lo tanto, la medida de cohesión depende del número de métodos, del número de atributos y de la relación entre ellos. La medida de acoplamiento está en relación al número de métodos y del número de dependencias de la clase.

Para este caso de prueba el mejor valor de las métricas de cohesión se da en el escenario (#Esce.2), este valor se obtiene cuando los atributos son usados por todos los métodos de la clase. En otro caso, cuando los atributos no son usados por todos los métodos de la clase se

obtiene el peor valor de cohesión (#Esce.12). La métrica CoupT presenta un decremento en su valor cuando el número de métodos en la clase aumenta y los canales de comunicación disminuyen (#Esce.11).

Tabla 5. Mediciones CP. V.A. Discretas

CP. Discretas																	
#Esce	Clase	#Met	#Atrib	ΣAxM	LCOM*		LCOM4		LCOM5		Coh		ClassCoh		CBO	CoupT	
					0 bueno 1 malo	1 bueno >1 malo	0 bueno 1 malo	1 bueno 0 malo	0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo	0 bueno 1 malo	0 bajo 1 alto				
1	E1_Discretas	12	3	11	0.76	1	0.71	0.31	0.465	0	0.00						
2	E2_Discretas	11	3	11	0.73	1	0.69	0.33	0.558	1	0.50						
3	E3_Discretas	10	3	11	0.70	1	0.66	0.37	0.681	1	0.60						
4	E4_Discretas	10	4	16	0.67	1	0.62	0.40	0.506	1	0.52						
5	E5_Discretas	9	4	14	0.69	1	0.63	0.39	0.472	2	0.57						
6	E6_Discretas	8	4	12	0.71	1	0.65	0.38	0.438	2	0.60						
7	E7_Discretas	7	2	8	0.50	1	0.46	0.57	0.524	3	0.65						
8	E8_Discretas	6	2	6	0.60	1	0.55	0.50	0.533	3	0.65						
9	E9_Discretas	4	2	4	0.67	1	0.57	0.50	0.333	3	0.76						
10	E10_Discretas	3	1	2	0.50	1	0.50	0.67	0.333	3	0.82						
11	E11_Discretas	2	1	1	1.00	1	1.00	0.50	0.000	2	0.83						
12	E12_Discretas	2	1	2	0.00	1	0.00	1.00	1.000	3	0.90						

En la tabla 5, se muestran las mediciones a nivel de clase de las métricas de cohesión y acoplamiento para el caso de prueba Discretas. En los escenarios se varió el número de métodos de la clase, provocando que la medida de cohesión y acoplamiento cambiaran.

Para este caso de prueba el mejor valor de las métricas de cohesión se da en el escenario (#Esce.12), este valor se obtiene cuando los atributos son usados por todos los métodos de la clase. Cuando los atributos son usados por solo un método de la clase se obtiene el peor valor de cohesión (#Esce.11). Para la métrica Coh el peor valor de cohesión (#Esce.1) se obtiene cuando el número de atributos usados por los métodos de la clase (ΣAxM) es menor a la mitad del total de atributos por método. Y para la métrica CoupT su mejor valor es cuando la clase tiene mayor número de métodos y menor número de canales de comunicación.

Tabla 6. Mediciones CP. Hiper cubo

CP. Hiper cubo														
#Esce	Clase	#Met	#Atrib	ΣAxM	LCOM*		LCOM4		LCOM5		Coh		CBO	CoupT
					0 bueno 1 malo	1 bueno >1 malo	0 bueno 1 malo	1 bueno >1 malo	0 bueno 0 malo	1 bueno 0 malo	0 bajo 1 alto			
1	E1_Hiper cubo	12	5	18	0.76	1	0.71	0.30	0.187	0	0.00			
2	E2_Hiper cubo	11	5	16	0.78	1	0.72	0.29	0.175	1	0.15			
3	E3_Hiper cubo	10	5	15	0.78	1	0.71	0.30	0.142	2	0.38			
4	E4_Hiper cubo	9	5	14	0.77	1	0.70	0.31	0.117	3	0.47			
5	E5_Hiper cubo	8	6	16	0.76	1	0.68	0.33	0.131	4	0.62			
6	E6_Hiper cubo	7	6	16	0.72	1	0.63	0.38	0.175	4	0.71			
7	E7_Hiper cubo	5	6	16	0.58	1	0.48	0.53	0.367	4	0.82			
8	E8_Hiper cubo	4	3	9	0.33	1	0.27	0.75	0.583	5	0.85			
9	E9_Hiper cubo	3	3	6	0.50	1	0.38	0.67	0.500	5	0.87			
10	E10_Hiper cubo	2	3	5	0.33	1	0.20	0.83	0.667	6	0.92			
11	E11_Hiper cubo	2	2	4	0.00	1	0.00	1.00	1.000	5	0.83			
12	E12_Hiper cubo	2	2	4	0.00	1	0.00	1.00	1.000	1	0.50			

En la tabla 6, se muestran las mediciones a nivel de clase de las métricas de cohesión y de acoplamiento para el caso de prueba Hiper cubo. Como se mencionó en los casos de prueba anteriores, el mejor y peor valor de las métricas de cohesión dependen del número de atributos que comparten los métodos de la clase.

En los escenarios (#Esce.11 y 12) de la tabla 6, la medida de cohesión no varió, pero la medida de acoplamiento sí, es decir, el cambio en el número de canales de comunicación no afecta para la medida de cohesión, debido a que no hay un incremento en los métodos de la clase principal (E#_Hiper cubo). Este es un problema cuando se tienen mediciones a nivel de clase, ya que solo se ve la cohesión de la clase principal, mas no la cohesión de las clases acopladas.

Tabla 7. Mediciones CP. TaylorBucket

CP. TaylorBucket																	
#Esce	Clase	#Met	#Atrib	$\Sigma A \times M$	LCOM*		LCOM4		LCOM5		Coh		ClassCoh		CBO	CoupT	
					0 bueno 1 malo	1 bueno >1 malo	0 bueno 1 malo	1 bueno 0 malo	0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo	0 bajo 1 alto					
1	E1_TaylorBucket	2	1	2	0.00	1	0.00	1.00	1.000	2	0.67						
2	E2_TaylorBucket	3	1	2	0.50	1	0.50	0.67	0.333	7	0.89						
3	E3_TaylorBucket	4	1	2	0.67	1	0.67	0.50	0.167	6	0.85						
4	E4_TaylorBucket	5	1	2	0.75	1	0.75	0.40	0.100	5	0.80						
5	E5_TaylorBucket	6	1	2	0.80	1	0.80	0.33	0.067	4	0.75						
6	E6_TaylorBucket	7	1	2	0.83	1	0.83	0.29	0.048	3	0.70						
7	E7_TaylorBucket	8	1	2	0.86	1	0.86	0.25	0.036	2	0.56						
8	E8_TaylorBucket	8	1	2	0.86	1	0.86	0.25	0.036	2	0.56						
9	E9_TaylorBucket	9	1	2	0.88	1	0.88	0.22	0.028	2	0.53						
10	E10_TaylorBucket	10	1	2	0.89	1	0.89	0.20	0.022	1	0.44						
11	E11_TaylorBucket	11	1	2	0.90	1	0.90	0.18	0.018	1	0.35						
12	E12_TaylorBucket	12	1	2	0.91	1	0.91	0.17	0.015	1	0.25						
13	E13_TaylorBucket	13	1	2	0.92	1	0.92	0.15	0.013	1	0.13						
14	E14_TaylorBucket	14	1	2	0.92	1	0.92	0.14	0.011	0	0.00						

En la tabla 7, se muestran las mediciones a nivel de clase de las métricas de cohesión y acoplamiento para el caso de prueba TaylorBucket. En los escenarios solo se varió el número de métodos de la clase, para conocer el comportamiento de la medida de acoplamiento con respecto a la medida de cohesión.

Para este caso de prueba, se observa que las mediciones de las métricas LCOM* y LCOM5 son iguales, debido a que el número de atributos y el número de atributos usados por los métodos de la clase se mantienen. También, se observa como la medida de acoplamiento mejora cuando el número de métodos en la clase aumentan y disminuye el número de canales de comunicación.

- **Mediciones a nivel de sistema:** Para obtener la medida de cohesión a nivel de sistema, se suma la medida de cohesión de cada clase del sistema y se divide entre el número total de ellas. No se consideran las clases donde la medida de cohesión es invalida.

A continuación, se presentan las tablas con las mediciones a nivel de sistema de cada caso de prueba. Los mejores y peores valores de las métricas se dan bajo las mismas condiciones encontradas en las mediciones a nivel de clase.

Tabla 8. Mediciones a nivel de sistema CP. Bernoulli

CP. Bernoulli										
#Esce	Paquete	#Clases	LCOM*	LCOM4	LCOM5	Coh	WTCoh	CBO	COF	WTCoup
			0 bueno 1 malo		0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo		0 bajo 1 alto	0 bajo 1 alto
1	E1_Bernoulli	1	?	1.0	0.000	1.000	?	0	0.000	0.000
2	E2_Bernoulli	1	?	1.0	1.000	1.000	?	0	0.000	0.000
3	E3_Bernoulli	1	?	1.0	1.000	0.000	0.000	0	0.000	0.000
4	E4_Bernoulli	1	0.500	1.0	0.333	0.750	0.500	0	0.000	0.000
5	E5_Bernoulli	1	0.950	1.0	0.905	0.136	0.009	0	0.000	0.000
6	E6_Bernoulli	3	0.917	2.0	0.812	0.262	0.016	2	0.333	0.147
7	E7_Bernoulli	3	0.667	2.0	0.723	0.317	0.127	2	0.333	0.139
8	E8_Bernoulli	2	0.833	2.5	0.742	0.330	0.065	1	0.500	0.389
9	E9_Bernoulli	2	0.729	2.5	0.662	0.393	0.149	1	0.500	0.368
10	E10_Bernoulli	3	0.750	2.0	0.756	0.280	0.090	2	0.333	0.139
11	E11_Bernoulli	13	0.250	0.9	0.026	0.981	0.500	13	0.083	0.019
12	E12_Bernoulli	13	0.375	0.9	0.046	0.962	0.167	12	0.077	0.019

La tabla 8, muestra el mejor escenario para el caso de prueba Bernoulli (#Esce.11), donde hay alta cohesión y bajo acoplamiento. La métrica LCOM4 en este escenario presenta un valor de 0.9, el cual indica que en el sistema existe una o más clases con cero métodos, donde su medida de LCOM4 es cero. Cuando la medida de LCOM4 es mayor a uno, significa que existen clases en el sistema con más de una responsabilidad.

Tabla 9. Mediciones del CP. Continuas

CP. Continuas										
#Esce	Paquete	#Clases	LCOM*	LCOM4	LCOM5	Coh	WTCoh	CBO	COF	WTCoup
			0 bueno 1 malo		0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo		0 bajo 1 alto	0 bajo 1 alto
1	E1_Continuas	20	?	1.0	0.150	0.850	?	21	0.055	0.020
2	E2_Continuas	20	0.000	1.0	0.150	0.850	1.000	21	0.055	0.020
3	E3_Continuas	19	0.500	1.0	0.185	0.825	0.333	20	0.058	0.022
4	E4_Continuas	19	0.500	1.0	0.184	0.825	0.333	20	0.058	0.020
5	E5_Continuas	18	0.667	1.0	0.204	0.806	0.167	19	0.062	0.022
6	E6_Continuas	18	0.833	1.0	0.206	0.799	0.083	19	0.062	0.022
7	E7_Continuas	19	0.875	1.0	0.191	0.811	0.167	20	0.058	0.020
8	E8_Continuas	18	0.889	1.0	0.207	0.796	0.167	19	0.062	0.022
9	E9_Continuas	17	0.950	1.0	0.223	0.779	0.100	17	0.062	0.022
10	E10_Continuas	16	0.971	1.0	0.239	0.762	0.067	15	0.062	0.021
11	E11_Continuas	13	0.976	1.0	0.296	0.705	0.048	11	0.071	0.022
12	E12_Continuas	13	1.000	1.0	0.308	0.705	0.000	11	0.071	0.022

La tabla 9, muestra el mejor escenario para el caso de prueba Continuas (#Esce.2), donde hay mejor nivel de cohesión y de acoplamiento. La métrica LCOM4 presenta valores de 1.0, lo que indica que las clases del sistema solo tienen una única responsabilidad.

Tabla 10. Mediciones del CP. Discretas

CP. Discretas										
#Esce	Paquete	#Clases	LCOM*	LCOM4	LCOM5	Coh	WTCoh	CBO	COF	WTCoup
			0 bueno 1 malo		0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo		0 bajo 1 alto	0 bajo 1 alto
1	E1_Discretas	1	0.758	1.0	0.714	0.306	0.465	0	0.000	0.000
2	E2_Discretas	2	0.733	1.0	0.344	0.667	0.558	1	0.500	0.250
3	E3_Discretas	2	0.704	1.0	0.828	0.183	0.341	1	0.500	0.300
4	E4_Discretas	2	0.667	1.0	0.808	0.200	0.253	1	0.500	0.262
5	E5_Discretas	4	0.344	0.8	0.657	0.597	0.236	3	0.250	0.103
6	E6_Discretas	4	0.238	0.8	0.411	0.594	0.479	3	0.250	0.106
7	E7_Discretas	6	0.125	0.7	0.410	0.762	0.508	5	0.167	0.066
8	E8_Discretas	6	0.120	0.7	0.258	0.750	0.633	5	0.167	0.069
9	E9_Discretas	6	0.133	0.7	0.262	0.750	0.583	5	0.167	0.067
10	E10_Discretas	6	0.100	0.7	0.250	0.778	0.583	5	0.167	0.066
11	E11_Discretas	6	0.200	0.7	0.333	0.750	0.500	4	0.133	0.069
12	E12_Discretas	7	0.000	0.7	0.143	0.857	0.750	5	0.119	0.051

La tabla 10, muestra el mejor escenario para el caso de prueba Discretas (#Esce.12), donde hay alta cohesión y bajo acoplamiento. La métrica LCOM4 en este escenario presenta un valor de 0.7, el cual indica que en el sistema existe una o más clases con cero métodos, donde su medida de LCOM4 es cero.

Tabla 11. Mediciones del CP. Hipercubo

CP. Hipercubo										
#Esce	Paquete	#Clases	LCOM*	LCOM4	LCOM5	Coh	WTCoh	CBO	COF	WTCoup
			0 bueno 1 malo		0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo		0 bajo 1 alto	0 bajo 1 alto
1	E1_Hipercubo	1	0.764	1.0	0.712	0.300	0.187	0	0.000	0.000
2	E2_Hipercubo	2	0.780	1.0	0.361	0.645	0.175	1	0.500	0.077
3	E3_Hipercubo	3	0.778	1.0	0.238	0.767	0.142	2	0.333	0.062
4	E4_Hipercubo	4	0.775	1.0	0.176	0.828	0.117	3	0.250	0.039
5	E5_Hipercubo	5	0.762	1.0	0.136	0.867	0.131	4	0.200	0.031
6	E6_Hipercubo	5	0.722	1.2	0.327	0.676	0.087	5	0.250	0.060
7	E7_Hipercubo	5	0.583	1.2	0.297	0.707	0.183	4	0.200	0.041
8	E8_Hipercubo	7	0.333	1.3	0.325	0.774	0.292	6	0.143	0.039
9	E9_Hipercubo	9	0.250	1.2	0.375	0.693	0.167	10	0.139	0.042
10	E10_Hipercubo	10	0.167	1.2	0.320	0.740	0.222	11	0.122	0.034
11	E11_Hipercubo	11	0.000	1.2	0.364	0.776	0.333	12	0.109	0.036
12	E12_Hipercubo	11	0.000	1.2	0.364	0.818	0.333	13	0.118	0.041

La tabla 11, muestra dos escenarios para el caso de prueba Hipercubo, donde el valor de cohesión es alto y el valor de acoplamiento es bajo. En el primero (#Esce.5), con las métricas de cohesión LCOM5 y Coh y la métrica de acoplamiento WTCoup. El segundo (#Esce.11), con las métricas LCOM* y WTCoh y las métricas de acoplamiento COF y WTCoup. En el primer escenario la métrica LCOM4 = 1.0; las clases del sistema tienen una única

responsabilidad. En el segundo $LCOM4 = 1.2$; una o más clases del sistema tienen más de una responsabilidad

Tabla 12. Mediciones del CP. TaylorBucket

CP. TaylorBucket										
#Esce	Paquete	#Clases	LCOM*	LCOM4	LCOM5	Coh	WTCoh	CBO	COF	WTCoup
			0 bueno 1 malo		0 bueno 1 malo	1 bueno 0 malo	1 bueno 0 malo		0 bajo 1 alto	0 bajo 1 alto
1	E1_TaylorBucket	11	0.000	1.0	0.182	1.000	1.000	9	0.082	0.021
2	E2_TaylorBucket	10	0.500	1.0	0.150	0.967	0.333	8	0.089	0.017
3	E3_TaylorBucket	9	0.667	1.0	0.185	0.944	0.167	7	0.097	0.021
4	E4_TaylorBucket	8	0.750	1.0	0.219	0.925	0.100	6	0.107	0.026
5	E5_TaylorBucket	7	0.800	1.0	0.257	0.905	0.067	5	0.119	0.034
6	E6_TaylorBucket	6	0.833	1.0	0.306	0.881	0.048	4	0.133	0.045
7	E7_TaylorBucket	5	0.857	1.0	0.371	0.850	0.036	3	0.150	0.061
8	E8_TaylorBucket	5	0.857	1.0	0.371	0.850	0.036	3	0.150	0.061
9	E9_TaylorBucket	4	0.875	1.0	0.219	0.806	0.028	2	0.167	0.044
10	E10_TaylorBucket	2	0.889	1.0	0.444	0.600	0.022	1	0.500	0.222
11	E11_TaylorBucket	2	0.900	2.0	0.950	0.091	0.009	1	0.500	0.176
12	E12_TaylorBucket	2	0.909	1.5	0.955	0.083	0.008	1	0.500	0.125
13	E13_TaylorBucket	2	0.917	1.0	0.458	0.577	0.013	1	0.500	0.067
14	E14_TaylorBucket	1	0.923	1.0	0.923	0.143	0.011	0	0.000	0.000

La tabla 12, muestra dos escenarios para el caso de prueba TaylorBucket, donde el valor de cohesión es alto y el valor de acoplamiento es bajo. En el primero (#Esce.1), con las métricas de cohesión LCOM*, Coh y WTCoh y la métrica de acoplamiento COF. El segundo (#Esce.2), con las métricas LCOM5 y Coh y la métrica de acoplamiento WTCoup. En los dos escenarios la métrica $LCOM4 = 1.0$; las clases del sistema tienen una única responsabilidad.

Capítulo 6: Análisis Estadísticos e Interpretación de Resultados

6.1 Análisis estadístico

Para conocer el grado y la forma de relación que existe entre las métricas de cohesión y acoplamiento, se llevó a cabo el análisis *de correlación y de Regresión*, usando el *método de Mínimos Cuadrados* para obtener los coeficientes de la función. El análisis de correlación permite encontrar un valor numérico que indica el sentido y la fuerza de una relación lineal entre dos variables. El análisis de Regresión se usa para describir la forma de la relación entre estas variables y para derivar una función matemática óptima que modele esta relación (3.4). En la mayoría de las ocasiones se espera que la relación entre variables se considere en el marco de la linealidad asumiendo como válida la ecuación de la recta (*Regresión Lineal*). Sin embargo, existen situaciones en las que esto no se puede dar. Por lo tanto, es preciso buscar otras funciones que representen de forma más adecuada la relación, como lo puede ser un polinomio resultado de un análisis de *Regresión Polinómica*.

El análisis de los datos inicia con la representación gráfica de las mediciones de las métricas de cohesión y acoplamiento, mediante un diagrama de dispersión, donde el “eje x” es representado por las mediciones de las métricas de cohesión y el “eje y” por las mediciones de las métricas de acoplamiento. Posteriormente, se obtiene el valor del coeficiente de correlación de Pearson (r), que indica el grado de relación lineal entre dos variables, si el valor de $r \geq 0.8$ la relación se considera *buena* o *excelente*. Se utiliza el método de los mínimos cuadrados para ajustar la recta a los puntos y encontrar los coeficientes de la función, para obtener el modelo lineal que representa la relación. En otro caso, si los puntos no tienen una tendencia lineal, se busca la curva con mejor ajuste, mediante el método de los mínimos cuadrados para el caso polinomial. Se obtiene el modelo que representa la relación y se calcula el coeficiente de determinación (r^2) que expresa en porcentaje, qué tan bien el modelo matemático representa la relación entre la métrica de cohesión y la métrica de acoplamiento.

Para este estudio, el análisis de las mediciones de las métricas de cohesión y acoplamiento se dividió en dos partes; conjunto de mediciones a nivel de clase y conjunto de mediciones a nivel de sistema.

6.1.1 Análisis #1: Mediciones de cohesión y acoplamiento a nivel de clase

Para tener una idea del sentido y forma de relación que existe entre las métricas de cohesión y acoplamiento, se presentan los diagramas de dispersión de las métricas de cohesión por las métricas de acoplamiento a nivel de clase.



Figura 14. Mediciones a nivel de clase del CP. Bernoulli

La figura 14, muestra las mediciones de las métricas de cohesión con la métrica de acoplamiento CoupT para el caso de prueba “Bernoulli”. Se observa en la dispersión de los puntos que no existe una correlación significativa entre estas métricas. Para conocer el grado y sentido de relación, se obtiene el valor del coeficiente de correlación lineal (r), para cada par de métricas.

Por ejemplo, para obtener el grado de relación lineal entre las métricas LCOM* y CoupT, se obtienen las sumatorias de las mediciones de LCOM* (variable x) y de las mediciones de CoupT (variable y), como se muestra en la tabla 13. Se sustituyen los valores de las sumatorias en la ecuación (20) y se obtiene el valor de r .

Tabla 13. Sumatorias de LCOM* y CoupT

x (LCOM*)	y (CoupT)	x^2	$x*y$	y^2
0.50	0.00	0.25	0	0.00
0.95	0.00	0.90	0	0.00
1.00	0.88	1.00	0.88	0.77
0.50	0.83	0.25	0.415	0.69
0.83	0.78	0.69	0.6474	0.61
0.62	0.74	0.38	0.4588	0.55
0.67	0.83	0.45	0.5561	0.69
0.50	0.91	0.25	0.455	0.83
0.75	0.86	0.56	0.645	0.74
6.32	5.83	4.74	4.0573	4.88

$$r = \frac{9 * 4.0573 - 6.32 * 5.83}{\sqrt{[9 * 4.74 - 39.94][9 * 4.88 - 33.98]}}$$

$$r = \frac{36.51 - 36.84}{5.19}$$

$$r = -0.064$$

Este cálculo se realizó para cada par de mediciones de cohesión y de acoplamiento. El valor de r indica el grado de relación lineal entre estas mediciones. En la tabla 14, se muestran los valores de r para las mediciones de cohesión y de acoplamiento del caso de prueba “Bernoulli”.

Tabla 14. Indicador de r para cohesión y CoupT – CP. Bernoulli

Métricas		Coefficiente de correlación lineal
x	y	r
LCOM*	CoupT	-0.064
LCOM5	CoupT	-0.198
Coh	CoupT	-0.028
ClassCoh	CoupT	0.151

Los valores de r están por debajo de 0.5, lo que representa una relación lineal *mala* entre estas métricas.

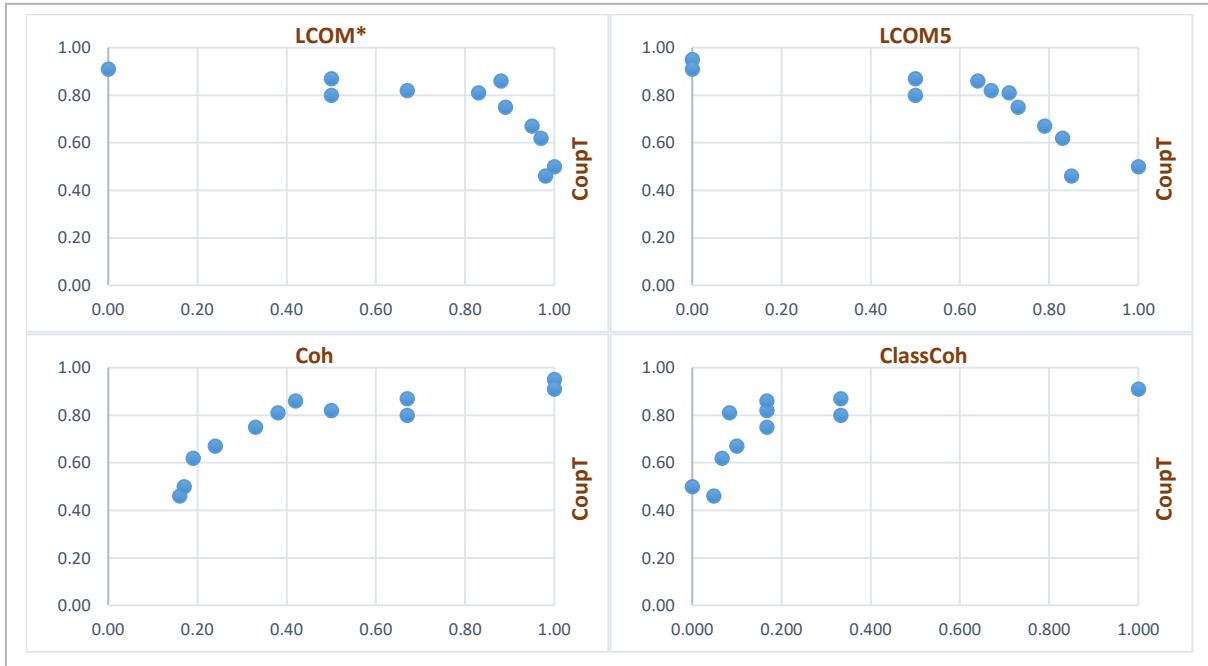


Figura 15. Mediciones a nivel de clase del CP. Continuas

La figura 15, muestra los diagramas de dispersión de las mediciones de cohesión y de acoplamiento (CoupT) del caso de prueba “Continuas”. Se observa en la dispersión de los puntos que existe una mejor correlación entre las medidas de cohesión y acoplamiento para este caso de prueba, en comparación con el caso “Bernoulli”. Para conocer el grado de relación lineal entre estas mediciones, se obtiene el valor del coeficiente de correlación (r), para cada par de métricas, de igual forma que el ejemplo anterior.

Tabla 15. Indicador de r para cohesión y CoupT - CP. Continuas

Métricas		Coefficiente de correlación lineal
x	y	r
LCOM*	CoupT	-0.697
LCOM5	CoupT	-0.798
Coh	CoupT	0.838
ClassCoh	CoupT	0.618

En la tabla 15, se muestran los valores de r para las mediciones de cohesión y acoplamiento del caso de prueba “Continuas”. Las métricas Coh y CoupT presentan una relación *buena*, las otras métricas de cohesión presentan una relación *regular* con la métrica CoupT. El sentido de la relación entre las mediciones de cohesión y de acoplamiento, en las métricas (LCOM* y LCOM5) es negativo o inverso a la métrica CoupT, es decir cuando la medida de una aumenta, la medida de la otra disminuye; considerando que estas métricas miden la carencia de cohesión, la relación se considera positiva, cuando una medida aumenta, la otra también aumenta.

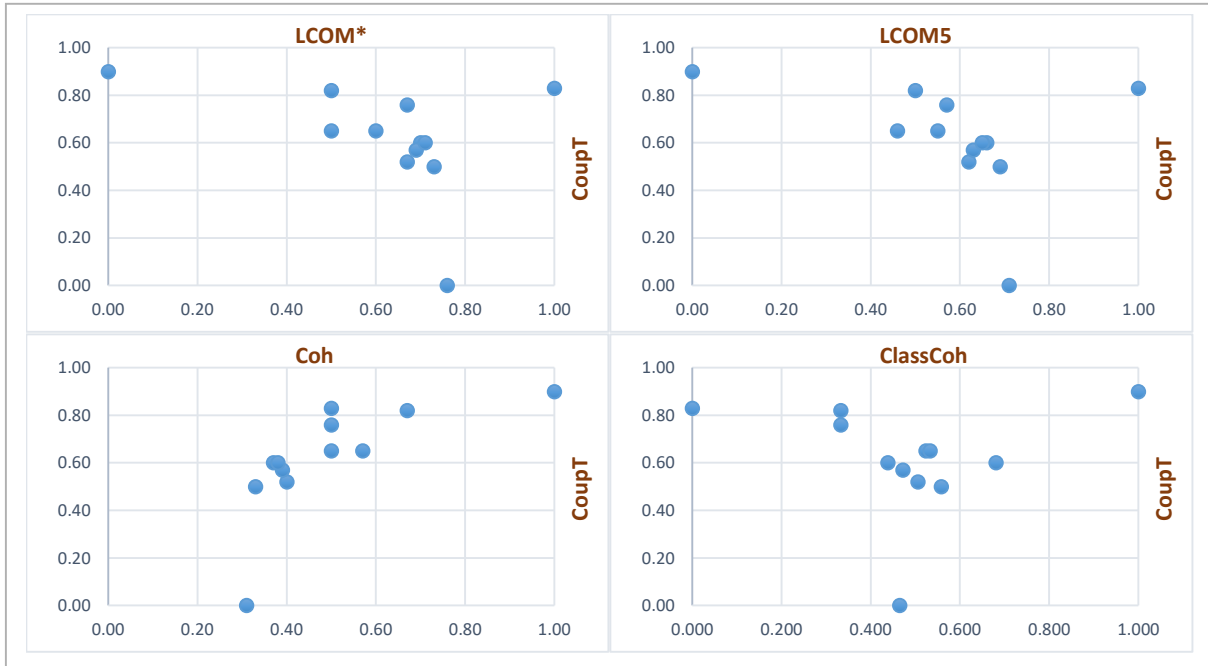


Figura 16. Mediciones a nivel de clase del CP. Discretas

La figura 16, muestra las mediciones de las métricas de cohesión con la métrica de acoplamiento CoupT del caso de prueba “Discretas” representadas en un diagrama de dispersión. Se observa en la dispersión de los puntos que no existe correlación significativa entre las métricas de cohesión y la métrica de acoplamiento (CoupT). Para conocer el grado de relación entre estas métricas, se obtiene el valor del coeficiente de correlación (r), para cada par de métricas.

Tabla 16. Indicador de r para cohesión y CoupT - CP. Discretas

Métricas		Coefficiente de correlación lineal
x	y	r
LCOM*	CoupT	-0.375
LCOM5	CoupT	-0.339
Coh	CoupT	0.684
ClassCoh	CoupT	-0.012

En la tabla 16, se muestran los valores de r para las mediciones de cohesión y acoplamiento del caso de prueba “Discretas”. Las métricas Coh y CoupT son las únicas que presentan una relación *regular*, las demás presentan una relación *mala*.

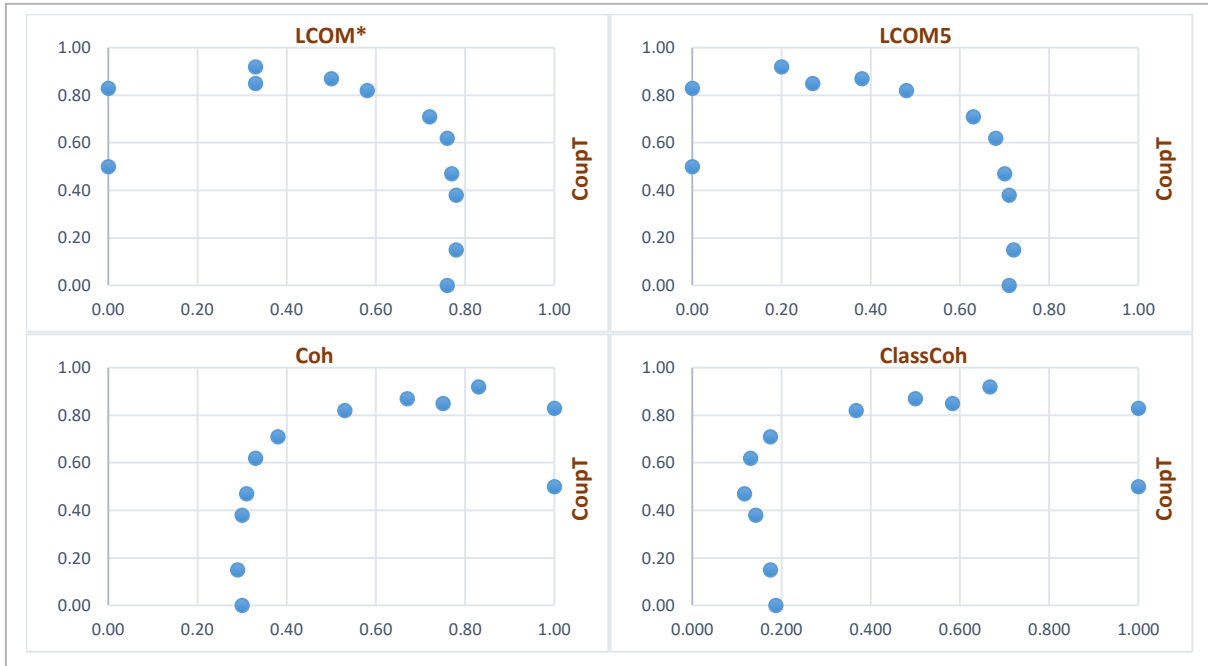


Figura 17. Mediciones a nivel de clase del CP. Hipercubo

La figura 17, muestra las mediciones de las métricas de cohesión con la métrica de acoplamiento CoupT del caso de prueba “Hipercubo”. Se observa que existe una correlación no lineal entre las métricas de cohesión y la métrica CoupT. Se obtiene el valor de r para comprobar si existe una relación lineal *buena* entre estas.

Tabla 17. Indicador de r para cohesión y CoupT - CP. Hipercubo

Métricas		Coefficiente de correlación lineal
x	y	r
LCOM*	CoupT	-0.502
LCOM5	CoupT	-0.583
Coh	CoupT	0.602
ClassCoh	CoupT	0.481

En la tabla 17, se muestran los valores de r para las mediciones de cohesión y acoplamiento del caso de prueba “Hipercubo”. Se observan valores bajos de r , indicando que la relación lineal no es *buena* entre estas métricas. Por lo tanto, las mediciones de este caso de prueba no deben tratarse en el marco de la linealidad.

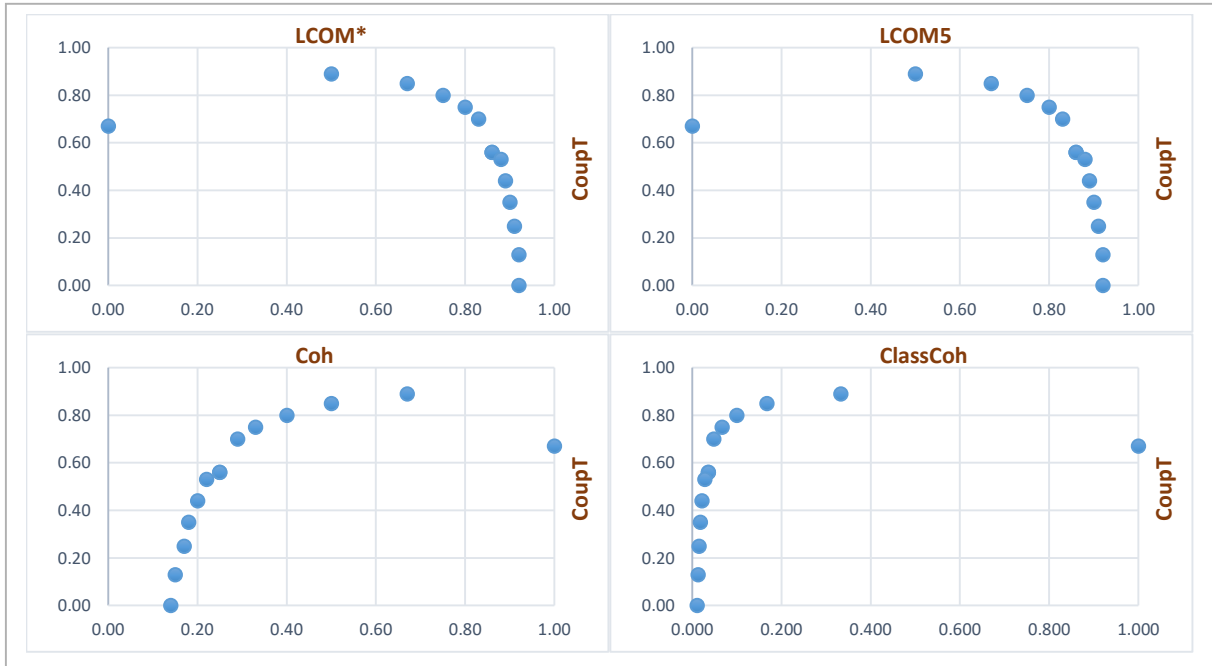


Figura 18. Mediciones a nivel de clase del CP. TaylorBucket

La figura 18, muestra las mediciones de las métricas de cohesión con la métrica de acoplamiento (CoupT) del caso de prueba “TaylorBucket”. Se observa en las gráficas que existe una relación curvilínea significativa entre estas métricas. Por lo tanto, este caso no se tratará en el marco de la linealidad.

De acuerdo a lo anterior, se buscó la curva que mejor se ajuste a las mediciones, es decir la función que mejor represente la relación entre las mediciones de cohesión y de acoplamiento. Para esto, se llevó a cabo un análisis de *Regresión Polinómica*, utilizando el método de los mínimos cuadrados para polinomios (18).

Considerando el sistema de ecuaciones para un polinomio de la forma $f(x) = a_0 + a_1x + a_2x^2$, se representa de la siguiente manera:

$$\begin{bmatrix} m & \sum x & \sum x^2 \\ \sum x & \sum x^2 & \sum x^3 \\ \sum x^2 & \sum x^3 & \sum x^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum yx \\ \sum yx^2 \end{bmatrix} \quad (22)$$

Para resolver el sistema de ecuaciones anterior, se utilizó el conjunto de mediciones de las métricas Coh (x) y CoupT (y) del caso de prueba de “TaylorBucket”. Se obtienen los valores

de las sumatorias requeridas en el conjunto de ecuaciones (tabla 18) y se sustituyen los valores. Por lo tanto, las ecuaciones simultaneas son:

Tabla 18. Sumatorias de Coh y CoupT

x Coh)	y (CoupT)	x ²	x ³	x ⁴	x*y	x ² *y
0.67	0.89	0.45	0.30	0.20	0.60	0.40
0.50	0.85	0.25	0.13	0.06	0.43	0.21
0.40	0.80	0.16	0.06	0.03	0.32	0.13
0.33	0.75	0.11	0.04	0.01	0.25	0.08
0.29	0.70	0.08	0.02	0.01	0.20	0.06
0.25	0.56	0.06	0.02	0.00	0.14	0.04
0.25	0.56	0.06	0.02	0.00	0.14	0.04
0.22	0.53	0.05	0.01	0.00	0.12	0.03
0.20	0.44	0.04	0.01	0.00	0.09	0.02
0.18	0.35	0.03	0.01	0.00	0.06	0.01
0.17	0.25	0.03	0.00	0.00	0.04	0.01
0.15	0.13	0.02	0.00	0.00	0.02	0.00
0.14	0.00	0.02	0.00	0.00	0.00	0.00
3.75	6.81	1.37	0.62	0.32	2.40	1.02

$$\begin{bmatrix} 13 & 3.75 & 1.37 \\ 3.75 & 1.37 & 0.62 \\ 1.37 & 0.62 & 0.32 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 6.81 \\ 2.40 \\ 1.02 \end{bmatrix}$$

Se resuelve el sistema de ecuaciones por el método de eliminación de Gauss (Pena, 2014) y se obtiene:

$$a_0 = -0.5687, \quad a_1 = 5.8091, \quad a_2 = -5.541$$

Se sustituyen los valores en la función polinómica (cuadrática) (14). Por lo tanto, el modelo matemático que representa la relación de cohesión (Coh) y acoplamiento (CoupT), se define de la siguiente manera:

$$y = -5.541x^2 + 5.8099x - 0.5687 \tag{23}$$

Usando las reglas de derivación, obtenemos la derivada de la función que representa la relación:

$$y = -5.541x^2 + 5.8099x - 0.5687$$

$$\frac{dy}{dx} = -11.082x + 5.8099$$

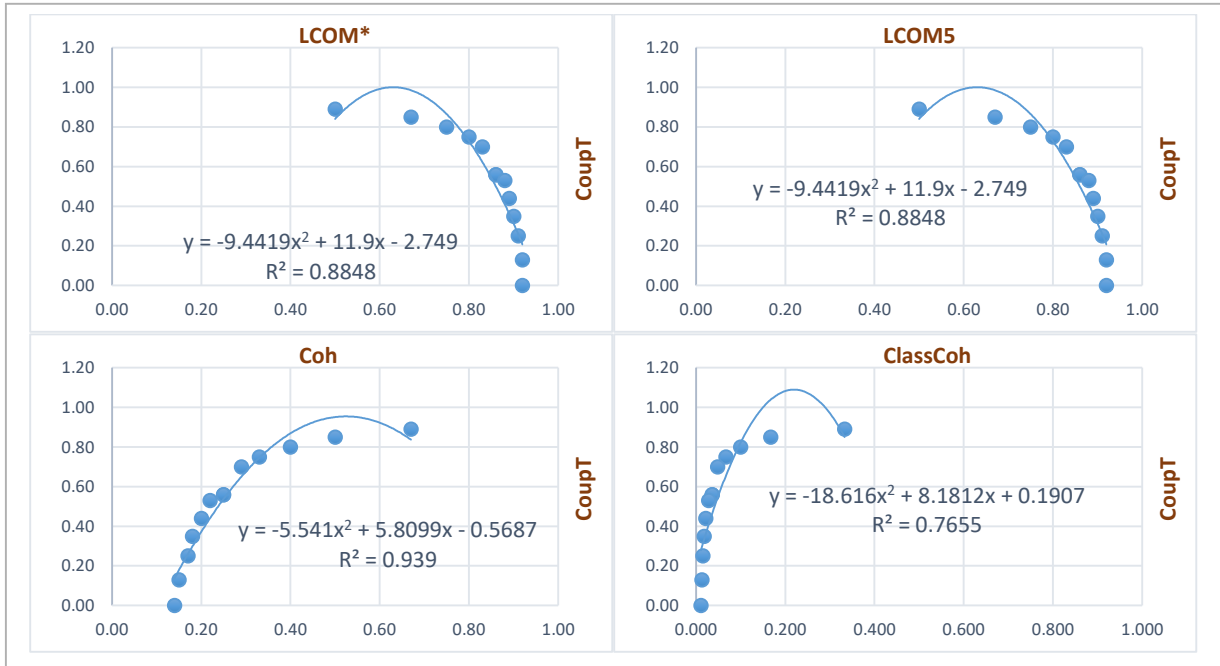


Figura 19. Función polinómica y r^2 de cohesión y CoupT – CP. TaylorBucket

En la figura 19 se muestra la función polinómica y el coeficiente de determinación (r^2) para las métricas de cohesión y la métrica CoupT del caso de prueba “TaylorBucket”.

Para calcular el coeficiente de determinación (r^2), se utiliza el modelo resultante (ecuación polinómica) (23), para obtener los valores de \hat{y} (mediciones estimadas de acoplamiento). Posteriormente se realizan algunos cálculos para obtener el error entre las mediciones reales y las mediciones estimadas de acoplamiento (ver tabla 19). Y finalmente se sustituyen las sumatorias de los errores en la formula (21) del coeficiente de determinación, como se muestra a continuación:

Tabla 19. Sumatorias de los errores al cuadrado - Coh y CoupT

X (Coh)	Y (CoupT)	\hat{Y}	SCR $= (\hat{Y} - \bar{Y})^2$	SCE $= (Y - \bar{Y})^2$
0.67	0.89	0.8365	0.0978	0.1340
0.50	0.85	0.9509	0.1824	0.1063
0.40	0.80	0.8686	0.1189	0.0762
0.33	0.75	0.7451	0.0489	0.0511
0.29	0.70	0.6501	0.0159	0.0310
0.25	0.56	0.5374	0.0002	0.0013
0.25	0.56	0.5374	0.0001	0.0013
0.22	0.53	0.4412	0.0068	0.0000
0.20	0.44	0.3716	0.0232	0.0070
0.18	0.35	0.2975	0.0512	0.0302
0.17	0.25	0.2588	0.0702	0.0750
0.15	0.13	0.1781	0.1195	0.1551
0.14	0.00	0.1360	0.1504	0.2744
	$\bar{Y} = 0.52$		$\Sigma 0.8855$	$\Sigma 0.9429$

$$r^2 = \frac{0.8855}{0.9429}$$

$$r^2 = 0.939$$

Otra forma de obtener (r) , es calculando la raíz cuadrada de (r^2) .

$$r = 0.969$$

El valor de r^2 , indica el porcentaje de predicción del modelo, en este caso el polinomial. Es decir que tanto la variable independiente (cohesión) explica a la variable dependiente (acoplamiento). El modelo que representa la relación entre Coh y CoupT tiene un 97% de predicción.

Estos mismos cálculos se realizaron para todas las métricas de cohesión y la métrica CoupT de los casos de prueba de TaylorBucket, Hipercubo y Continuas, dado que sus graficas muestran una tendencia no lineal.

La tabla 20, muestra el grado de relación entre las métricas de cohesión y la métrica de acoplamiento CoupT en el caso de prueba TaylorBucket. Dando como resultado que las métricas de cohesión y la métrica CoupT, tienen una relación *excelente* entre ellas. Excepto las métricas ClassCoh y CoupT se considera una relación *buena*.

Tabla 20. Indicador de r para cohesión y CoupT - CP. TaylorBucket

Métricas		Regresión polinómica
x	y	r
LCOM*	CoupT	-0.941
LCOM5	CoupT	-0.941
Coh	CoupT	0.969
ClassCoh	CoupT	0.875

En la figura 20, se muestra la función polinómica y el coeficiente de determinación (r^2) para las métricas de cohesión y la métrica CoupT del caso de prueba “Hipercubo”.

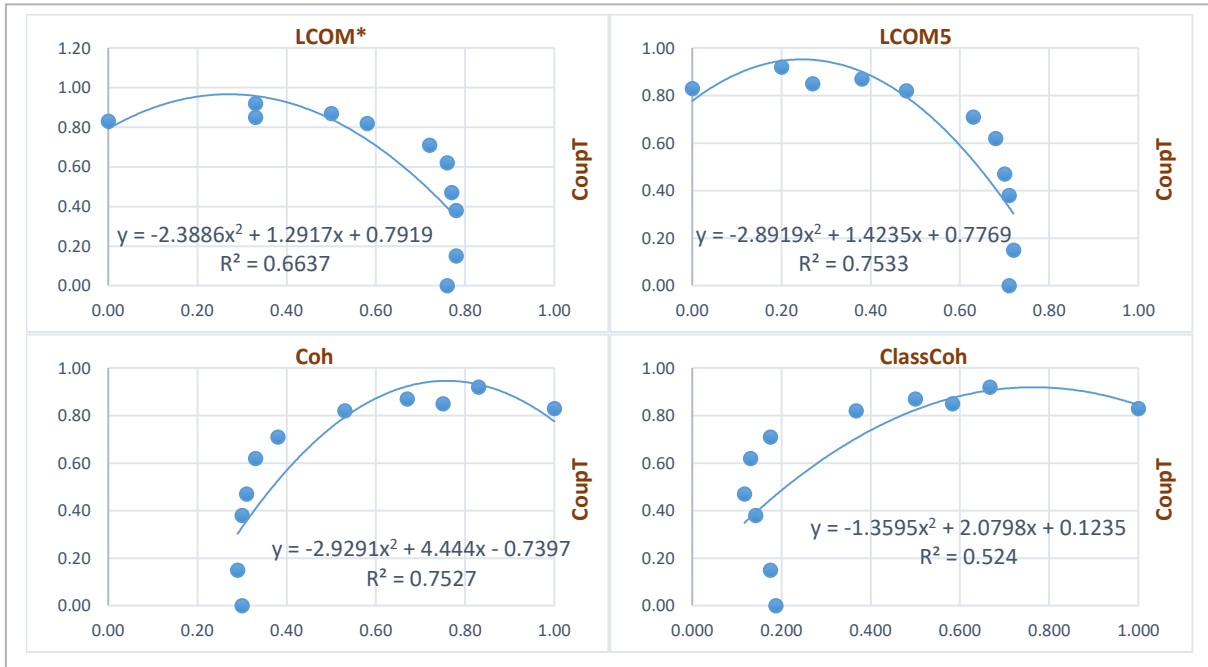


Figura 20. Función polinómica y r^2 de cohesión y CoupT – CP.Hipercubo

La tabla 21, muestra los valores de los coeficientes de correlación para la regresión polinómica. Dando como resultado que las métricas de cohesión y la métrica CoupT, tienen una relación *buena* entre ellas. Excepto las métricas ClassCoh y CoupT se considera una relación *regular*.

Tabla 21. Indicador de r para cohesión y CoupT - CP. Hipercubo

Métricas		Regresión polinómica
x	y	r
LCOM*	CoupT	-0.815
LCOM5	CoupT	-0.868
Coh	CoupT	0.867
ClassCoh	CoupT	0.723

En la figura 21, se muestra la función polinómica y el coeficiente de determinación (r^2) para las métricas de cohesión y la métrica CoupT del caso de prueba “Continuas”.

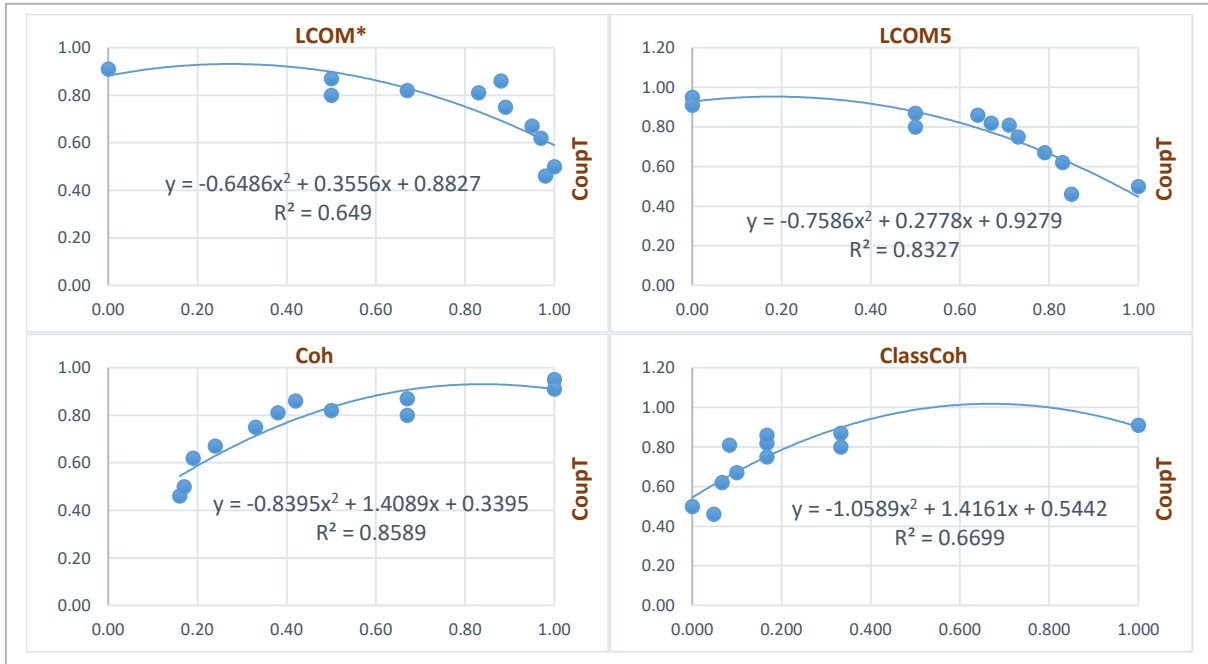


Figura 21. Función polinómica y r^2 de cohesión y CoupT – CP. Continuas

La tabla 22, muestra los valores de los coeficientes de correlación para la regresión polinómica. Dando como resultado que las métricas de cohesión (LCOM* y ClassCoh) y la métrica (CoupT), tienen una relación *buena* entre ellas. Las métricas (LCOM5 y Coh) y la métrica (CoupT) presentan una relación *excelente*.

Tabla 22. Indicador de r para cohesión y CoupT - CP. Continuas

Métricas		Regresión polinómica
x	y	r
LCOM*	CoupT	-0.806
LCOM5	CoupT	-0.913
Coh	CoupT	0.927
ClassCoh	CoupT	0.818

6.1.2 Análisis #2: Mediciones de cohesión y acoplamiento a nivel de sistema

Para este estudio, se consideran las mediciones a nivel de sistema de las métricas de cohesión (LCOM*, LCOM5, WTCoh, Coh) y las métricas de acoplamiento (COF y WTCoup) para cada caso de prueba. De la misma manera que el análisis #1, se obtuvieron los valores de los coeficientes (r) y (r^2) y se encontró la función que mejor representa la relación entre estas métricas.

Tabla 23. Indicador de r para cohesión con acoplamiento - CP. Bernoulli

Métricas		Regresión lineal	Regresión polinómica
x	y	R	r
LCOM*	COF	0.448	0.603
LCOM5	COF	0.247	0.737
Coh	COF	-0.416	-0.802
WTCoh	COF	-0.322	-0.610
LCOM*	WTCoup	0.433	0.565
LCOM5	WTCoup	0.246	0.684
Coh	WTCoup	-0.374	-0.737
WTCoh	WTCoup	-0.204	-0.535

En la tabla 23, se muestran los valores del coeficiente de correlación, tanto para la regresión lineal como para la regresión polinómica, de las mediciones de cohesión y de acoplamiento del caso de prueba “Bernoulli”. Los valores de r son más altos en la regresión polinómica, por lo tanto, la relación entre las métricas de cohesión y de acoplamiento tiene la forma polinómica y no lineal. Las métricas Coh y COF presentan una relación *buena*, mientras que las demás métricas de cohesión presentan una relación *regular* con las métricas de acoplamiento COF y WTCoup. Cuando el valor de r es negativo la relación es inversa, es decir cuando el valor de cohesión aumenta, el valor de acoplamiento disminuye, esto se puede ver en las gráficas.

A continuación, en la figura 22 y 23, se muestra la función polinómica y el coeficiente de determinación (r^2) para las métricas de cohesión y las métricas de acoplamiento COF y WTCoup del caso de prueba “Bernoulli”.

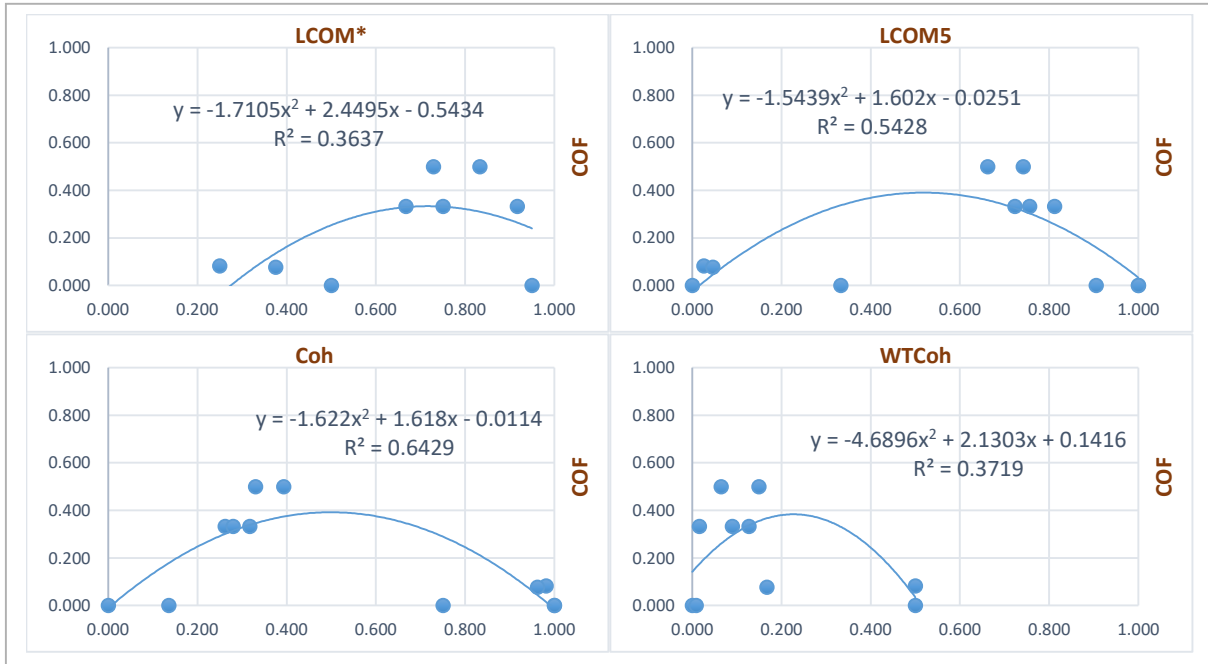


Figura 22. Función polinómica y r^2 de cohesión y COF – CP. Bernoulli

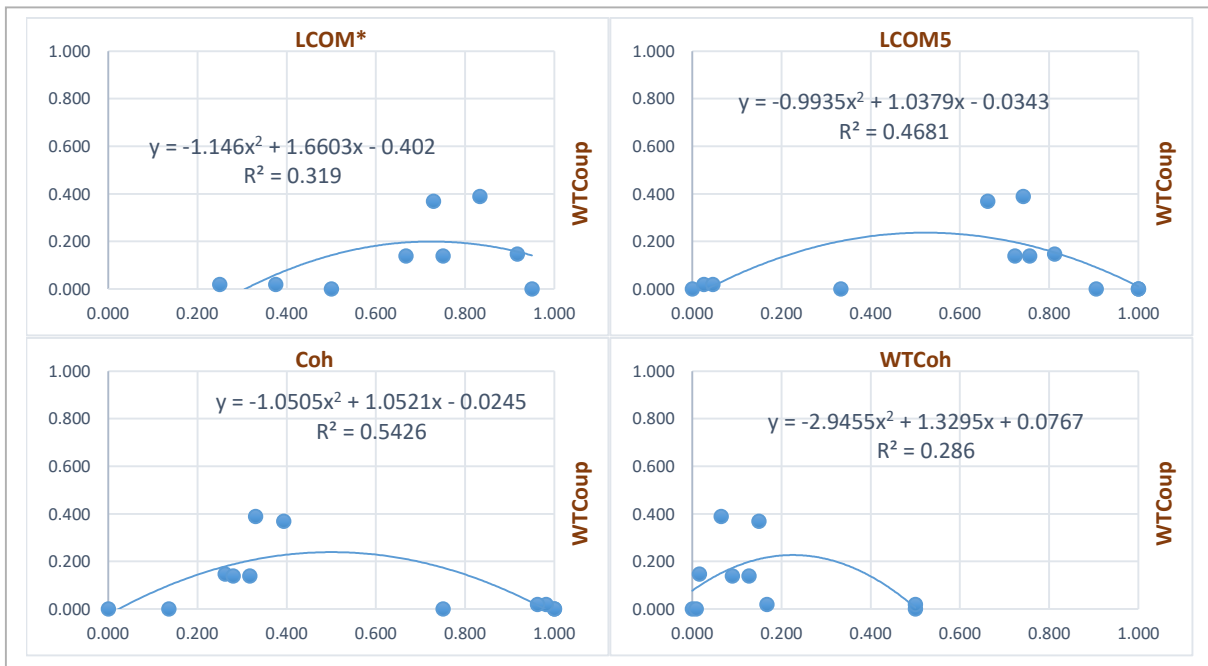


Figura 23. Función polinómica y r^2 de cohesión y WTCoup – CP. Bernoulli

Tabla 24. Indicador de r para cohesión con acoplamiento - CP. Continuas

Métricas		Regresión lineal	Regresión polinómica
x	y	r	r
LCOM*	COF	0.708	0.754
LCOM5	COF	0.978	0.978
Coh	COF	-0.972	-0.972
WTCoh	COF	-0.682	-0.854
LCOM*	WTCoup	0.520	0.536
LCOM5	WTCoup	0.606	0.709
Coh	WTCoup	-0.583	-0.682
WTCoh	WTCoup	-0.582	-0.596

En la tabla 24, se muestran los valores de r , para la regresión lineal y para la regresión polinómica, de las mediciones de las métricas de cohesión y las métricas de acoplamiento del caso de prueba “Continuas”. En algunos casos los valores de r son un poco más altos en la regresión polinómica, sin embargo, esa variación no es significativa, por lo tanto, la relación entre las métricas de cohesión y de acoplamiento puede considerarse como lineal para el caso de prueba “Continuas”. Las métricas LCOM5 y COF, Coh y COF presentan una relación *excelente*, mientras que las de más métricas de cohesión presentan una relación *regular* con las métricas de acoplamiento COF y WTCoup. Las métricas LCOM5 y COF, tienen una relación positiva o directa, es decir cuando el valor de carencia de cohesión aumenta, también aumenta el valor de acoplamiento.

A continuación, en la figura 24 y 25, se muestra la función lineal y el coeficiente de determinación (r^2) para las métricas de cohesión y las métricas de acoplamiento COF y WTCoup del caso de prueba “Continuas”.

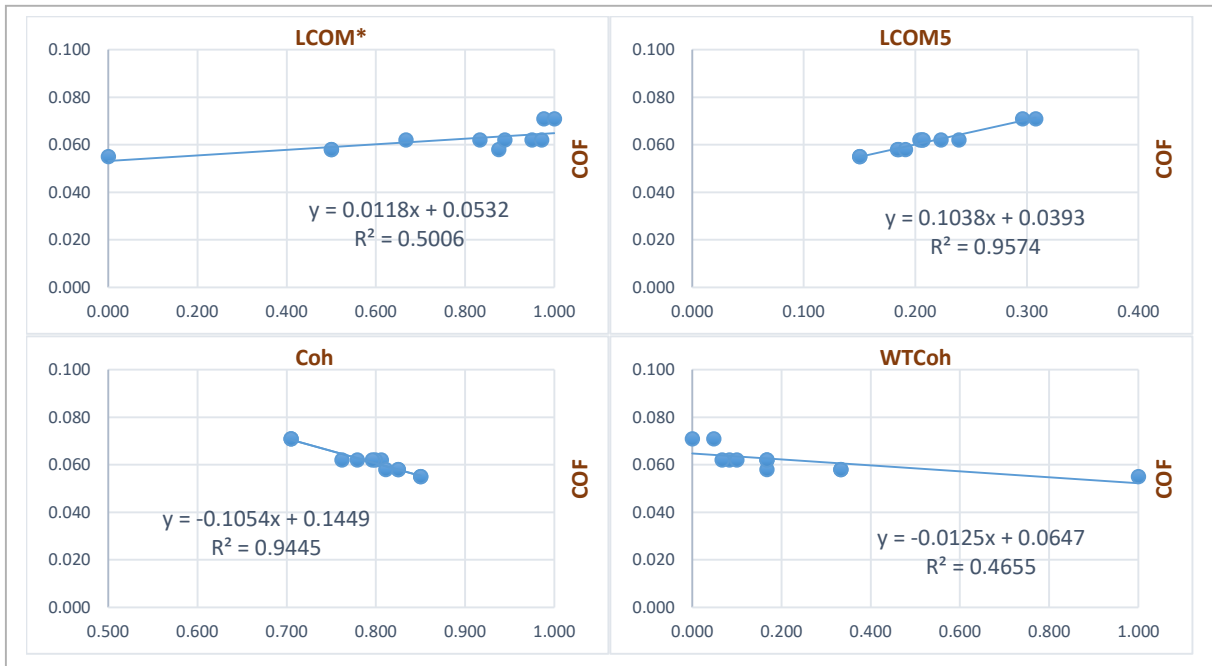


Figura 24. Función lineal y r^2 de cohesión y COF – CP. Continuas

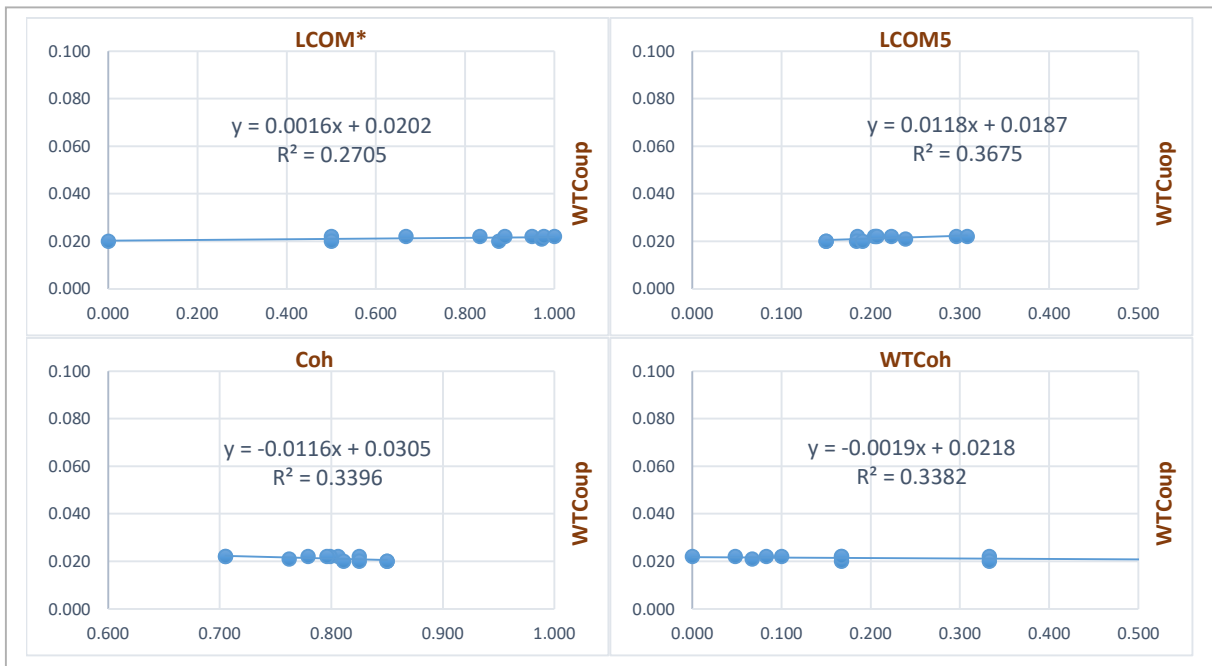


Figura 25. Función lineal y r^2 de cohesión y WTCoup – CP. Continuas

Tabla 25. Indicador de r de cohesión con acoplamiento - CP. Discretas

Métricas		Regresión lineal	Regresión polinómica
x	y	r	r
LCOM*	COF	0.569	0.600
LCOM5	COF	0.452	0.476
Coh	COF	-0.501	-0.503
WTCoh	COF	-0.491	-0.500
LCOM*	WTCoup	0.627	0.640
LCOM5	WTCoup	0.517	0.564
Coh	WTCoup	-0.584	-0.599
WTCoh	WTCoup	-0.500	-0.505

En la tabla 25, se muestran los valores de r , para la regresión lineal y para la regresión polinómica, de las mediciones de cohesión y de acoplamiento del caso de prueba “Discretas”. Los valores de r son un poco más altos en la regresión polinómica, sin embargo, la diferencia no es significativa, por lo tanto, la relación entre las métricas de cohesión y de acoplamiento puede considerarse como lineal para el caso de prueba “Discretas”. Cuando el valor se encuentra $0.5 \leq r \leq 0.8$ la relación entre las métricas se considera *regular*, y cuando $r < 0.5$ la relación es *mala*.

A continuación, en la figura 26 y 27, se muestra la función lineal y el coeficiente de determinación (r^2) para las métricas de cohesión y las métricas de acoplamiento COF y WTCoup del caso de prueba “Discretas”.

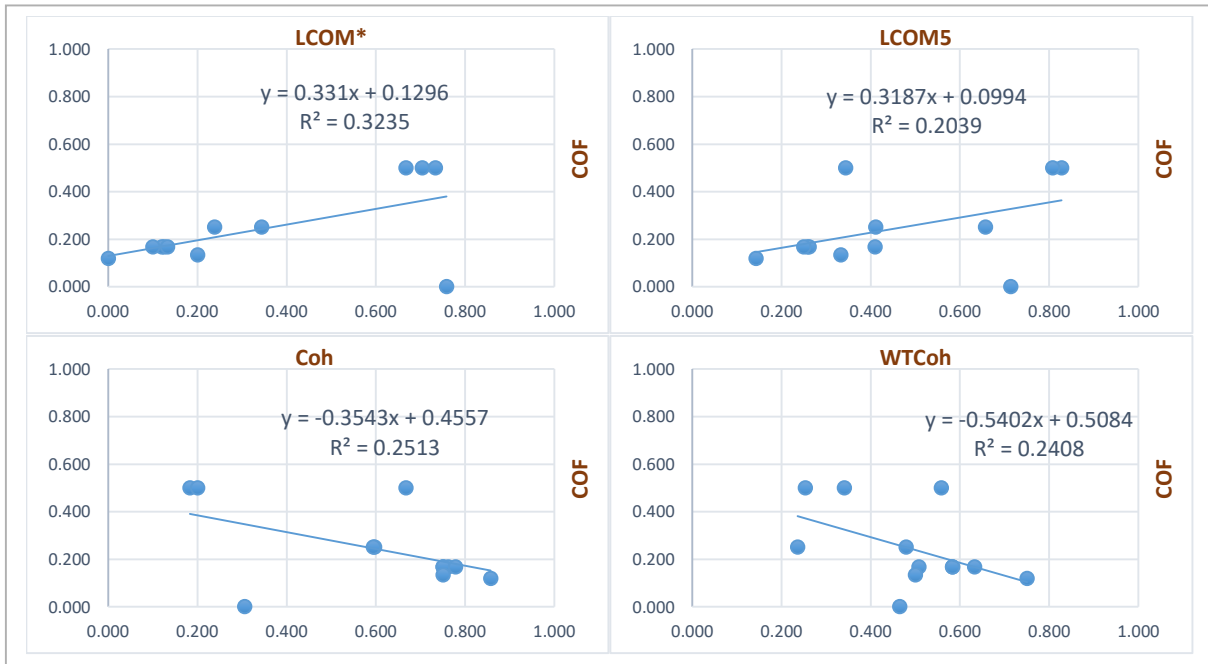


Figura 26. Función lineal y r^2 de cohesión y COF – CP. Discretas

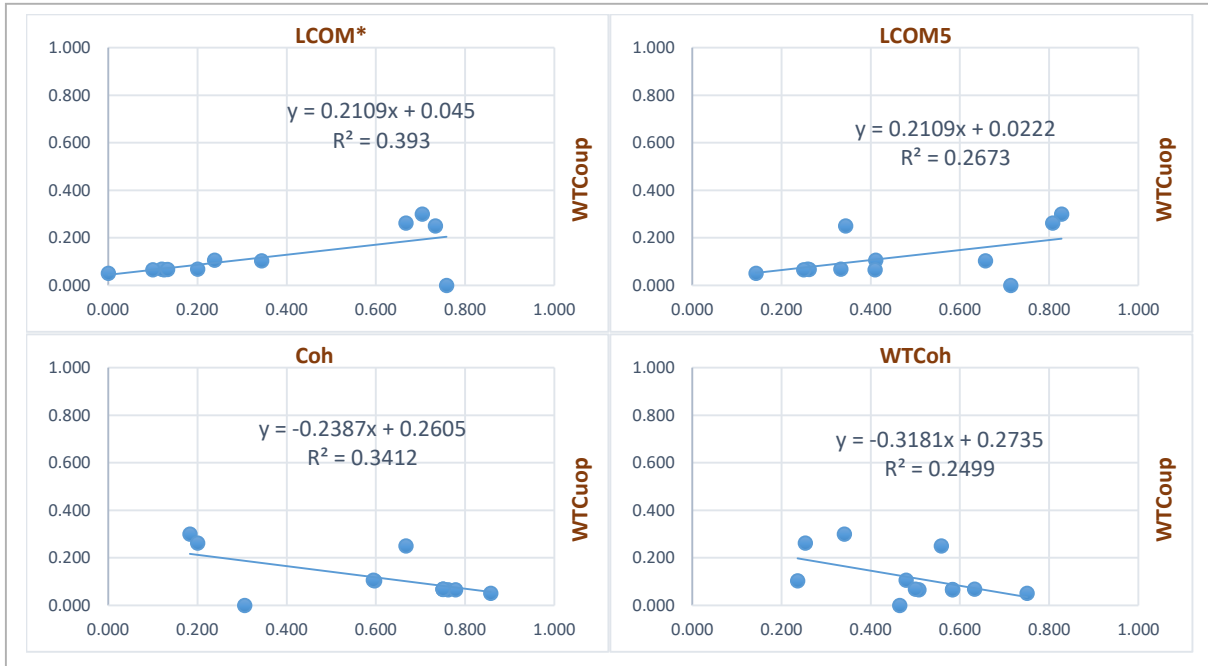


Figura 27. Función lineal y r^2 de cohesión y WTCoup – CP. Discretas

Tabla 26. Indicador de r de cohesión con acoplamiento - CP. Hiper cubo

Métricas		Regresión lineal	Regresión polinómica
x	y	r	r
LCOM*	COF	0.510	0.526
LCOM5	COF	-0.472	-0.503
Coh	COF	0.265	0.538
WTCoh	COF	-0.436	-0.437
LCOM*	WTCoup	0.167	0.173
LCOM5	WTCoup	-0.485	-0.732
Coh	WTCoup	0.408	0.803
WTCoh	WTCoup	-0.227	-0.315

En la tabla 26, se muestran los valores de r , para la regresión lineal y para la regresión polinómica, de las mediciones de cohesión y de acoplamiento del caso de prueba “Hiper cubo”. En las métricas LCOM5 y WTCoup, Coh y WTCoup, los valores de r de la regresión lineal varían significativamente con respecto a r de la regresión polinómica, por lo tanto, la relación entre las métricas de cohesión y de acoplamiento se considera como polinómica para el caso de prueba “Hiper cubo”.

A continuación, en la figura 28 y 29, se muestra la función polinómica y el coeficiente de determinación (r^2) para las métricas de cohesión y las métricas de acoplamiento COF y WTCoup del caso de prueba “Hiper cubo”.

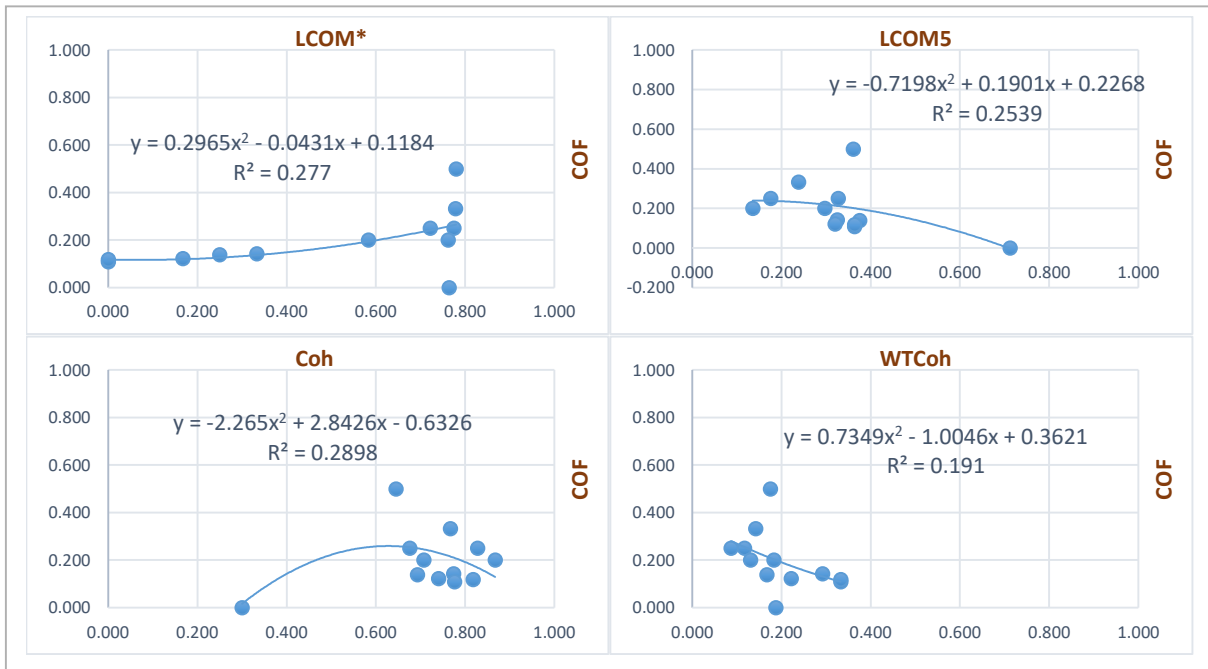


Figura 28. Función polinómica y r^2 de cohesión y COF – CP. Hipercubo

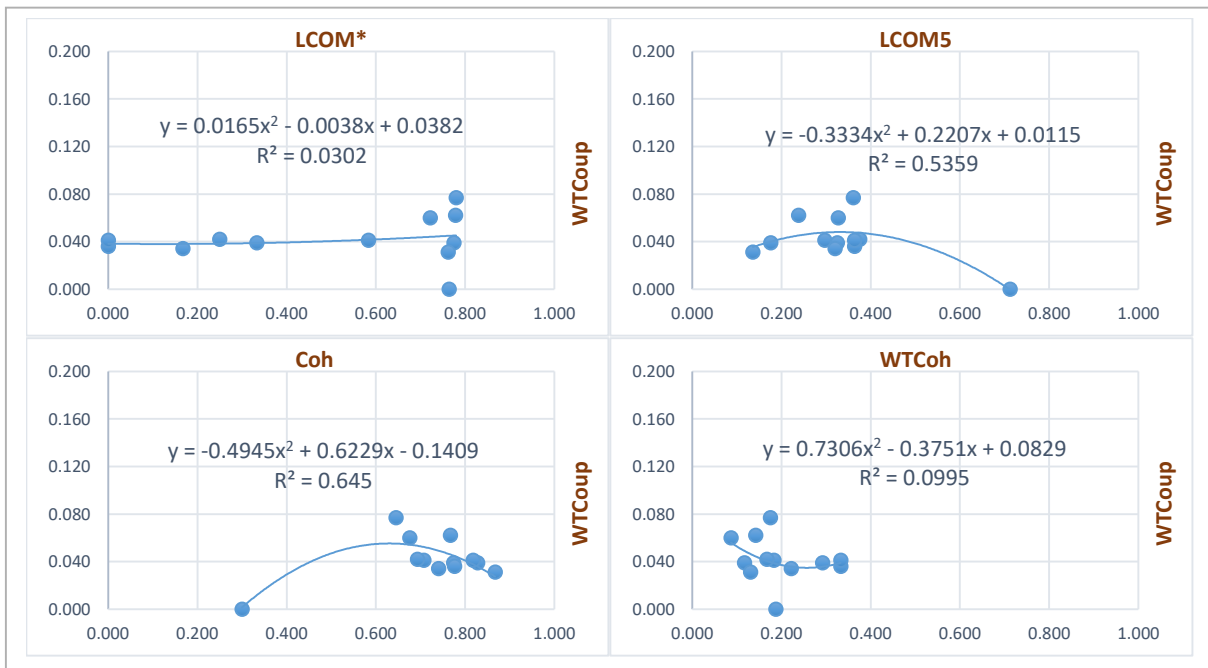


Figura 29. Función polinómica y r^2 de cohesión y WTCoup – CP. Hipercubo

Tabla 27. Indicador de r de cohesión con acoplamiento - CP. TylorBucket

Métricas		Regresión lineal	Regresión polinómica
x	y	r	r
LCOM*	COF	0.407	0.554
LCOM5	COF	0.513	0.621
Coh	COF	-0.580	-0.703
WTCoh	COF	-0.343	-0.510
LCOM*	WTCoup	0.378	0.516
LCOM5	WTCoup	0.466	0.589
Coh	WTCoup	-0.504	-0.617
WTCoh	WTCoup	-0.420	-0.486

En la tabla 27, se muestran los valores de r , para la regresión lineal y para la regresión polinómica, de las mediciones de cohesión y de acoplamiento del caso de prueba “TylorBucket”. Los valores de r de la regresión lineal varían con respecto a los de la regresión polinómica, algunas variaciones son significativas y otras no, sin embargo, la relación entre las métricas de cohesión y de acoplamiento se considera polinómica para el caso de prueba “TylorBucket”. Ninguna de las relaciones entre estas métricas se considera *buena*, debido a que el valor de r está por debajo de 0.8.

A continuación, en la figura 30 y 31, se muestra la función polinómica y el coeficiente de determinación (r^2) para las métricas de cohesión y las métricas de acoplamiento COF y WTCoup del caso de prueba “TylorBucket”.

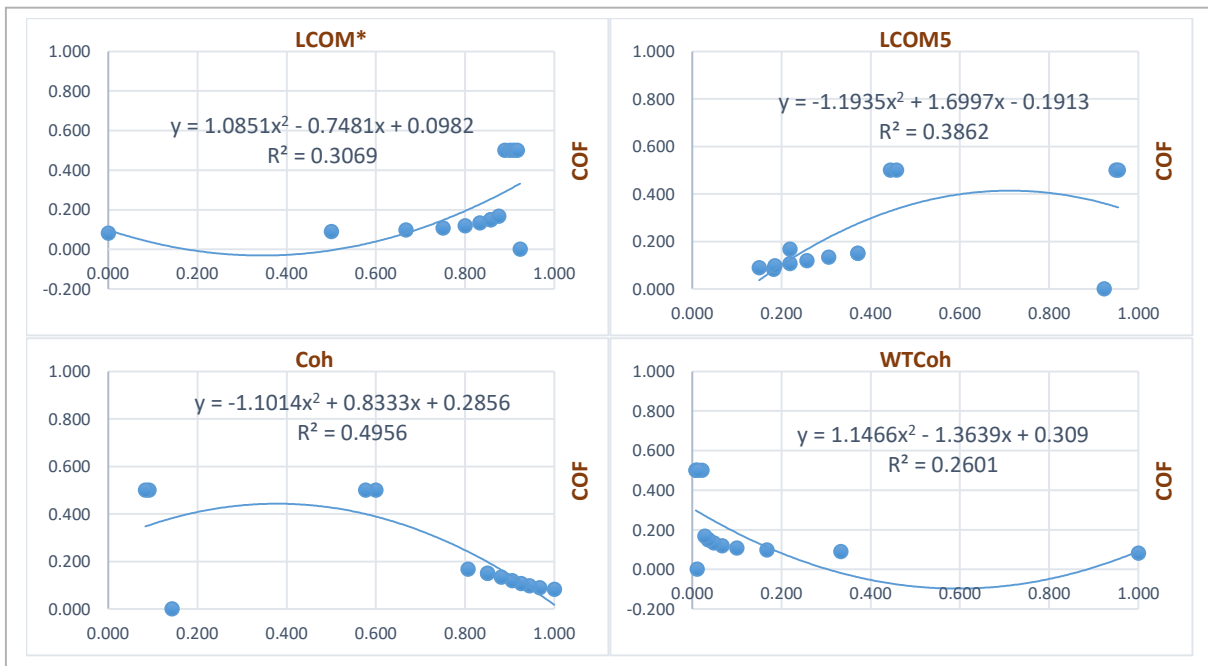


Figura 30. Función polinómica y r^2 de cohesión y COF – CP. TylorBucket

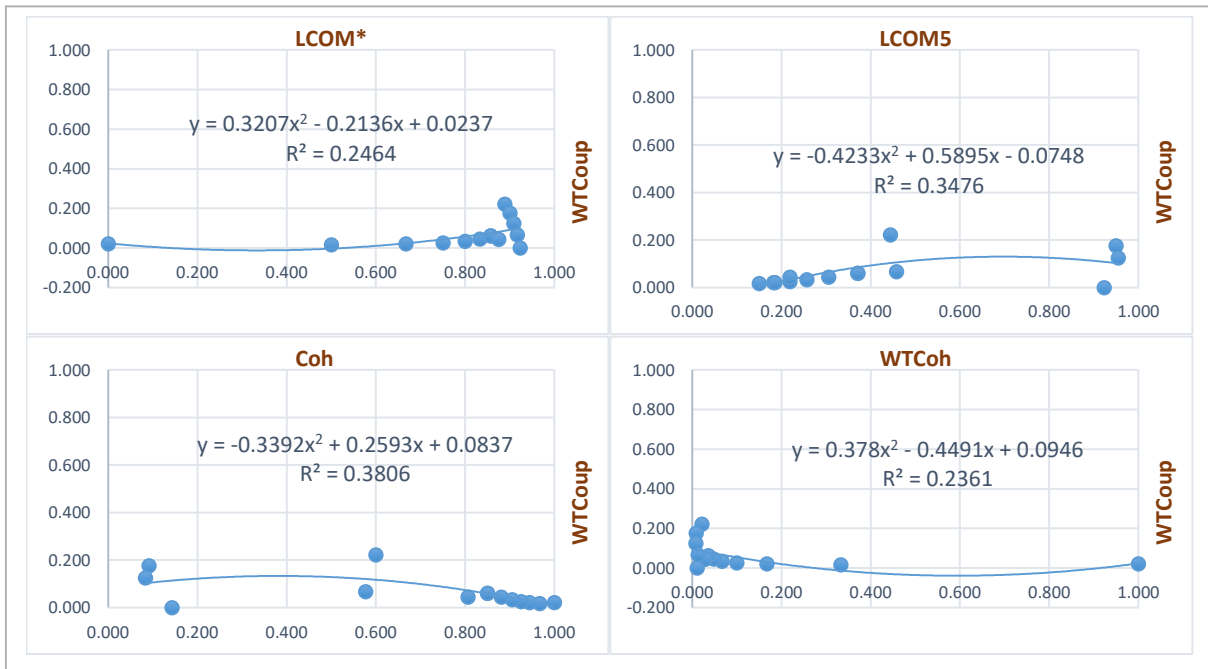


Figura 31. Función polinómica y r^2 de cohesión y WT Coup – CP. TylorBucket

Capítulo 7: Conclusiones y Trabajos Futuros

7.1 Conclusiones

Los resultados obtenidos mediante el análisis estadístico, dieron respuesta a la pregunta de investigación de esta tesis, ¿Cuál es la influencia y el grado de relación que tienen entre sí las métricas de cohesión y acoplamiento? El grado de relación de las métricas depende principalmente del nivel de medición, es decir, a nivel de clase o a nivel de sistema, así como de los parámetros que éstas tienen en común, como el número de métodos en la clase.

Las mediciones a nivel de clase del caso de prueba “TaylorBucket” presentan una relación *excelente*, las métricas Coh y CoupT con un grado de relación de 0.969 y LCOM* y CoupT con 0.941. También, las mediciones a nivel de sistema del caso de prueba “Continuas” presentan una relación *excelente* con las métricas LCOM5 y COF con un grado de relación de 0.978, así como Coh y COF con 0.972. En conclusión, se rechaza la hipótesis nula H_0 , la cual enuncia que “el grado de influencia o relación de las métricas de cohesión y de acoplamiento es nulo en unidades de programa”.

Las métricas de cohesión y acoplamiento a nivel de clase, presentan una relación positiva o directa, es decir, cuando la variable independiente (cohesión) aumenta, la variable dependiente (acoplamiento) también aumenta, aunque no en la misma proporción. Para el caso de las métricas de carencia de cohesión, la relación es inversa, cuando el valor de carencia de cohesión aumenta, el valor de acoplamiento disminuye. En ninguno de los dos casos es el ideal (alta cohesión – bajo acoplamiento). Esto se debe a que la medida de cohesión considera solo una clase, sin importar la medida de cohesión de las clases con la que esta acoplada. Sin embargo, la medida de acoplamiento si considera el conjunto de clases con la que la clase principal esta acoplada, por lo tanto, se considera que las métricas no están en el mismo nivel de medición. Las mediciones a nivel de clase cumplen con la hipótesis estadística H_2 planteada en esta investigación, la cual enuncia que “el grado de relación de la medida de acoplamiento es directamente proporcional a la medida de cohesión”.

Las métricas de cohesión y acoplamiento a nivel de sistema, presentan una relación negativa o inversa, es decir, cuando la variable independiente (cohesión) aumenta, la variable dependiente (acoplamiento) disminuye, aunque no en la misma proporción. Esto se da por que las métricas se encuentran en el mismo nivel de medición, es decir, la medida de cohesión y de acoplamiento consideran el total de las clases del sistema. Por lo tanto, las mediciones a nivel de sistema cumplen con la hipótesis estadística H_1 planteada en esta investigación, la cual enuncia que “el grado de relación de la medida de acoplamiento es inversamente proporcional a la medida de cohesión”.

La relación de las métricas de cohesión y de acoplamiento a nivel de clase tienen una forma no lineal, es decir, la función que representa la relación entre estas métricas tiene la forma polinómica de grado dos (cuadrática). La relación entre estas métricas no es perfecta, debido a que las mediciones de cohesión y de acoplamiento (x , y) no aparecen sobre la función, por lo que existe un mínimo de error al ajustar la función. La función polinómica con mejor ajuste y, por lo tanto, con el mínimo de error, es la que representa la relación entre las métricas de Coh y CoupT del caso de prueba “TylorBucket”, con un porcentaje de predicción (r^2) de 93%. Con esto se cumple el objetivo específico de representar la relación mediante un modelo matemático (función matemática).

Para las mediciones a nivel de sistema, las métricas LCOM5 y COF del caso de prueba “Continuas” presentan una relación lineal con un porcentaje de predicción del 95%. Esta forma se debe a que el rango de valores de acoplamiento es muy pequeño, debido al diseño de los escenarios de los casos de prueba, el cual se enfocó en la manipulación de la clase principal y no en el conjunto de clases del sistema.

Los resultados obtenidos cumplen con el objetivo principal de esta tesis, “obtener conocimiento acerca de la relación existente entre las métricas de cohesión y acoplamiento”.

7.2 Trabajos futuros

Los resultados obtenidos en esta investigación, muestran la influencia sobre el sentido de la relación de las métricas de cohesión y acoplamiento, de conformidad con el diseño de los escenarios de los casos de prueba. Como trabajo futuro, se propone validar la influencia que tienen las métricas de cohesión y acoplamiento ante escenarios en los se considere el diseño de todas las clases del sistema y no solo de la clase principal. Así mismo incrementar el número de canales de comunicación para que la medida de acoplamiento a nivel de sistema sea mayor.

La autonomía se define en razón a los atributos de cohesión y acoplamiento. A partir de las funciones que representan la relación de los atributos de cohesión y acoplamiento a nivel de clase y a nivel de sistema, obtenidas en este trabajo de tesis, se propone un trabajo de investigación para determinar, si es posible, una ecuación diferencial que represente el atributo de autonomía. Posterior a este trabajo, se propone realizar un estudio de optimización para determinar el valor ideal de cohesión y de acoplamiento para alcanzar el mejor valor de la autonomía.

Referencias Bibliográficas

- [1]. Abreu, J. L. (2012). Hipótesis, Método & Diseño de Investigación. *Daena: International Journal of Good Conscience*, 7(2), 187–197.
- [2]. Abreu y Melo. (1996). Evaluating the impact of ObjectOriented Design on Software Quality. *Proceedings of 3rd International Software Metrics Symp., Berlin*.
- [3]. Al Dallal, J. (2011). Measuring the discriminative power of object-oriented class cohesion metrics. *IEEE Transactions on Software Engineering*, 37(6), 788–804. <https://doi.org/10.1109/TSE.2010.97>
- [4]. Al Dallal, J. (2013). Object-oriented class maintainability prediction using internal quality attributes. *Information and Software Technology*, 55(11), 2028–2048. <https://doi.org/10.1016/j.infsof.2013.07.005>
- [5]. Alejandro Rosas Snell, J. Z. C. (2013). Estadística Descriptiva E Inferencial I. Retrieved from https://www.conevyt.org.mx/bachillerato/material_bachilleres/cb6/5sempdf/edin1/edin1_f03.pdf
- [6]. Beck, F., & Diehl, S. (2011). On the congruence of modularity and code coupling. *Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering - SIGSOFT/FSE '11*, 354. <https://doi.org/10.1145/2025113.2025162>
- [7]. Bidve, V. S., & Sarasu, P. (2016). Coupling measures and its impact on object-oriented software quality. *Indian Journal of Science and Technology*, 9(21). <https://doi.org/10.17485/ijst/2016/v9i21/89912>
- [8]. Bivde, V. S., & Sarasu, P. (2016). Correlation Between Coupling Metrics Values and Number of Classes in Multimedia Java Projects: A Case Study. *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(1), 78. <https://doi.org/10.9781/ijimai.2016.4114>
- [9]. Briand LC, Daly J, W. J. (1998). A unified framework for cohesion measurement in object-oriented systems.
- [10]. Chidamber, S. R., & Kemerer, C. F. (1994). A Metrics Suite for Object Oriented Design. *IEEE Transactions on Software Engineering*, 20(6), 476–493. <https://doi.org/10.1109/32.295895>
- [11]. Dirac, P. A. M. (1984). Modelos matemáticos.
- [12]. Elaine, J. and V. F. I. (2000). *Experience with performance testing of software systems: Issues, approach and case study*. (I. transactions on software Engineering, Ed.) (Vol. 26, N).
- [13]. Escudero, P. J. V. (2001). Métricas Orientadas a Objetos, (May 2014).
- [14]. Fisher, R. A. (1935). *The Design of Experiments*.
- [15]. Gui, G., & Scott, P. D. (2006). Coupling and cohesion measures for evaluation of component reusability. *Proceedings of the 2006 International Workshop on Mining Software Repositories - MSR '06*, 18. <https://doi.org/10.1145/1137983.1137989>
- [16]. Gui, Gui, & Scott, P. D. (2009). Measuring software component reusability by coupling and cohesion metrics. *Journal of Computers*, 4(9), 797–805. <https://doi.org/10.4304/jcp.4.9.797-805>

- [17]. Henderson-Sellers, B. (1996). Object-Oriented Metrics Measures of Complexity.
- [18]. Hernández Roberto, F. C. (2004). *Metodología de la Investigación*. (McGraw-Hill, Ed.). México D.F.
- [19]. Hitz, M., & Montazeri, B. (1995). Measuring coupling and cohesion in object-oriented systems. *Proceedings of the International Symposium on Applied Corporate Computing*, 50, 75–76. <https://doi.org/10.1.1.409.4862>
- [20]. IEEE. (1993). Software Engineering Standards, *Standard 6*.
- [21]. Kalantari, S., Alizadeh, M., & Motameni, H. (2015). Evaluation of reliability of object-oriented systems based on Cohesion and Coupling Fuzzy computing, 6(1), 85–99.
- [22]. Liskov, B. (1980). *A Design Methodology for Reliable Software Systems, in Tutorial on Software Design Techniques* (Third Edit). New York, NY.
- [23]. Mendiburu, F. de. (2007). Análisis de Regresión y Correlación, (X), 1–9.
- [24]. Orellana, L. (2008). *ANÁLISIS DE REGRESIÓN*.
- [25]. Paixao, M., Harman, M., Zhang, Y., & Yu, Y. (2017). An Empirical Study of Cohesion and Coupling: Balancing Optimisation and Disruption. *IEEE Transactions on Evolutionary Computation*, (c), 1–1. <https://doi.org/10.1109/TEVC.2017.2691281>
- [26]. Pena, M. G. y J. M. (2014). Interpolación , eliminación y matrices totalmente positivas Resumen, (January 1992).
- [27]. Pulido, H. G., & De la Vara, R. S. (2008). *Análisis y diseño de experimentos*. (McGRAW-HILL/INTERAMERICANA, Ed.).
- [28]. Ramón, J. M. (2012). No Title.
- [29]. Rathee, A., & Chhabra, J. K. (2018). Improving Cohesion of a Software System by Performing Usage Pattern Based Clustering. *Procedia Computer Science*, 125, 740–746. <https://doi.org/10.1016/j.procs.2017.12.095>
- [30]. Robledo, L. A. C. (2004). *Refactorización de Marcos Orientados a Objetos para Reducir el Acoplamiento Aplicando el Patrón de Diseño Mediator*.
- [31]. Rodríguez, D., & Harrison, R. (n.d.). Medición en la Orientación a Objetos. *Medición Para La Gestión En La Ingeniería de Software*, 1–16.
- [32]. Rogel, L. C. G. (2014). *Estudio Experimental para Determinar la Influencia del Nivel de Granulación de Servicios Web en Factores de Calidad QoS*. CENIDET.
- [33]. Roger S. Pressman. (2010). *Ingeniería de software - Un Enfoque practico*. *Danielr.Obolog.Es*.
<https://doi.org/http://zeus.inf.ucv.cl/~bcrawford/Modelado%20UML/Ingenieria%20del%20Software%207ma.%20Ed.%20-%20Ian%20Sommerville.pdf>
- [34]. Salomón Ortiz Quintanilla, Jorge Segura Davila, R. C. V. (2005). MÉTODOS NUMÉRICOS APLICADOS A LA INGENIERÍA. <https://doi.org/https://doi.org/10.33326/26176033.2005.9.174>
- [35]. Sánchez, A. M. G. (2010). *calidad del software sobre un programa Java . Programación y Tecnología Software*. <https://doi.org/10.7497/j.issn.2095-3941.2013.01.001>
- [36]. Sicilia, M. (2009). Métricas Orientadas a Objetos. *Management*, 2–4.
- [37]. Strub, A. M. (2015). Métricas estadísticas asociadas al proceso de desarrollo y mejora continua de software. *Congreso Virtual Sobre Tecnología, Educación y Sociedad*, 1(5).
- [38]. Yadav, S., Sikka, S., & Shrivastava, U. (2014). A review of Coupling and

Cohesion metrics in Object Oriented Environment. *International Journal of Computer Science & Engineering Technology (IJCSET)*, 2(5), 101–108.