Research Article

Method of Coupling Metrics for Object-Oriented Software System Based on CSBG Approach

Aihua Gu,1 Lu Li,1 Shujun Li,1 Qifeng Xun,1 Jian Dong,1 and Jianhong Lin2

1School of Information Engineering, Yancheng Teachers University, Yancheng 224002, China
2Zhejiang Ponshine Information Technology Co., Ltd., Hangzhou 310012, China

Correspondence should be addressed to Aihua Gu; guaihua1978@163.com

Received 2 November 2019; Accepted 5 February 2020; Published 19 March 2020

Guest Editor: Weifeng Pan

Copyright © 2020 Aihua Gu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Context. Coupling between classes is an important metric for software complexity in software systems. Objective. In order to overcome the shortcomings of the existing coupling methods and fully investigate the weighted coupling of classes in different cases in large-scale software systems, this study analyzed the relationship between classes at package level, class level, and method level. Method. The software system is considered as a set of special bipartite graphs in complex networks, and an effective method for coupling measurement is proposed as well. Furthermore, this method is theoretically proved to satisfy the mathematical properties of coupling measurement, leading to overcome the disadvantages of the majority of existing methods. In addition, it was revealed that the proposed method was efficient according to the analyses of existing methods for coupling measurement. Eventually, an algorithm was designed and a program was developed to calculate coupling between classes in three open-source software systems. Results. The results indicated the scale-free characteristic of complex networks in the statistical data. Additionally, the calculated power-law value was used as a metric for coupling measurement, so as to calculate coupling of the three open-source software. It indicated that coupling degrees of the open-source software systems contained a certain impact on evaluation of software complexity. Conclusions. It indicated that coupling degrees of the open-source software systems contained a certain impact on evaluation of software complexity. Moreover, statistical characteristics of some complex networks provided a reliable reference for further in-depth study of coupling. The empirical evidence showed that within a certain range, reducing the coupling was helpful to attenuate the complexity of the software, while excessively blindly pursuit of low coupling increases the complexity of software systems.

1. Introduction

Coupling refers to the degree of interdependence between software modules; a measure of how closely connected two routines or modules are [1]; and the strength of the relationships between modules. Structured design, including cohesion and coupling, was published in an article by Stevens et al. and a book by Stevens et al. [2, 3], and the latter subsequently became standard terms. Coupling is considered as a double-edged sword in object-oriented programming. On the one hand, object-oriented software development (OOSD) includes object-oriented requirement analysis, as well as object-oriented design. OOSD is a practical method of developing a software system which focuses on the objects of a problem throughout development. Interactions between objects reflect the interdependence between objects. If objects are isolated, then the software system can only achieve simple functions. However, objects are equivalent to cells in human body. If cells are completely isolated from human body, they basically do not play any significant role, reflecting that functions of a software system require a tight coupling between objects. On the other hand, tight coupling between objects would lead to a water-wave effect, meaning that changes in one object may result in further changes in other objects. The most terrible case is that there is a possibility of “avalanche” effect, which may affect the whole system, leading to a sharp decline in the testability, understandability, reliability, and maintainability.
of the system. Therefore, it is expected that classes are loosely coupled in terms of software design. A system can be tightly coupled in one aspect while being loosely coupled in another. However, software developers mainly prefer to develop those systems that are as loosely coupled as possible; thus, design, testing, and maintenance of the system would be relatively independent and more reasonable. Moreover, a decrease may be observed in the possibility that errors propagate between modules if there are few connections between modules [4]. Coupling has been widely used in evaluation of the degree of failure in classification [5–7], effective analysis [8, 9], and design pattern [10] of software systems.

The present article has the following organization: Section 2 summarizes the materials and methods. Section 3 shows the results. Section 4 provides a conclusion and suggests perspectives.

2. Materials and Methods

2.1. Methods. Currently, the methods for coupling measurement of object-oriented software systems are mainly structure-oriented measurement methods (Tables 1 and 2) [8, 11–15].

Comparative analysis of typical methods is shown in Tables 1–3, indicating that

1. Methods for coupling measurement are mainly based on method invocation between classes.
2. Calculation of coupling strength is defined as the degree of method invocation, which is weighted coupling.
3. A small number of methods use fan-in and fan-out as metrics.
4. Inheritance is dominant.
5. Few methods use static method invocation, system measurement, and package-level metrics.

In addition to the abovementioned structural information methods, some scholars have recently used dynamic information methods [17, 18], semantic information methods [19–21], and logical information methods [22, 23]. Based on the results of previous studies, methods of coupling measurement cover the following cases:

1. The DCMs are more accurate than that of structural information methods, while it seems to be difficult in the measurement of coupling metrics. However, structural information methods are more intuitive and easier to be perceived compared with semantic information and logical information methods.
2. At present, the majority of the traditional structural information methods analyze coupling based on the degree of connecting edges between classes and mainly focus on complexity between classes and emphasize more on measurement from a local fine-grained aspect. Moreover, these degrees of connecting edges only consider a certain or a limited aspects of software engineering. Therefore, these methods contain some limitations, which cannot properly satisfy the requirement of an effective coupling measurement in complex software systems.
3. Although a number of coupling measurement methods analyze network relationship from overall and macro perspectives based on graph theory, the majority of measurement metrics mainly use classes, packages, or methods as nodes to construct some undirected, directed, unweighted, or weighted network models. Moreover, they ignore a complex relationship of object-oriented characteristics between different classes. Some methods have not been theoretically verified for developing the mathematical characteristics of the measurement metrics.

The process of establishment of an effective method for coupling measurement between classes in a software system is determined by the following two aspects: reasonable measurement metrics and theoretical support of measurement metrics [24, 25]. Briand et al. mathematically analyzed measurement metrics of the software system and presented a robust theoretical support for the measurement metrics [4, 26, 27]. Many of complex networks have been shown to share the features such as “scale-free” and “small world” [28, 29]. Pan et al. revealed many physics-like laws in software systems from a complex network perspective recently [30, 31]. Studies on complex networks and software engineering revealed that class-level, method-level, and package-level diagrams of a software system could show the characteristics of “scale-free” and “small world,” which provided a novel perspective for finding more reasonable measurement metrics [32–34]. Complex network theories were applied to measure software [35, 36], identify key software elements [37], and cluster Web services [38–39]. Researchers have found that many real networks have the bipartite graph characteristics of complex networks [40–45]. With combination of package level, class level, and method level, this study analyzed a complex relationship between classes in the same layer and all layers of a package and proposed a method for coupling metrics for object-oriented systems based on bipartite graph of complex networks, named here CSBG, and object-oriented software systems were expressed as a set of special bipartite graphs.

2.2. Problem Description. In this study, a statistical method for complex networks was used to analyze the degree of fan-out and the heterogeneity of classes at the same layer and all layers of a package, in addition to the calculation of coupling of software systems.

2.2.1. Relationship between Classes

Definition 1. ASS relationship (association and aggregation).

Association means which/how classes interact with each other, and association can be represented by a line between these classes with an arrow indicating the navigation
**Table 1: The first part of existing methods for coupling measurement.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBO [11]</td>
<td>$CBO(c) =</td>
</tr>
<tr>
<td>CBO' [12]</td>
<td>$CBO'(c) =</td>
</tr>
<tr>
<td>RFC [13]</td>
<td>$RFC(c) = RFC_1(c)$, which is used for calculating the number of methods responding to an object’s message</td>
</tr>
<tr>
<td>RFCa [13]</td>
<td>$R_{Fc}(c) = M(c)$, $R_{Fc}(c) = \bigcup_{m \in \text{Rel}(c)} \text{PIM}(m)$, that is, a set of polymorphic methods invoked by functions in set $R_i(c)$; then $RFC_\alpha(c) =</td>
</tr>
<tr>
<td>RFC' [13]</td>
<td>$RFC'(c) = RFC_\alpha(c)$</td>
</tr>
<tr>
<td>MPC [13]</td>
<td>$MPC(c) = \sum_{m \in \text{Rel}(c)} \sum_{m' \in \text{SIM}(m)}</td>
</tr>
<tr>
<td>DAC [14]</td>
<td>$DAC(c) =</td>
</tr>
<tr>
<td>DAC' [14]</td>
<td>$DAC'(c) =</td>
</tr>
</tbody>
</table>

**Table 2: The second part of existing methods for coupling measurement.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>COF [14]</td>
<td>$COF(C) = \sum_{c \in \text{C}}</td>
</tr>
<tr>
<td>ICP [14]</td>
<td>This method calculates the parameters invoked by the method in a weighted class</td>
</tr>
<tr>
<td>IH-ICP [14]</td>
<td>This is similar to ICP, however, that only considers inheritance</td>
</tr>
<tr>
<td>NIH-ICP [14]</td>
<td>This is similar to ICP, however, that does not consider inheritance</td>
</tr>
<tr>
<td>SIMAS [8]</td>
<td>This method calculates the number of direct or indirect invocations between static methods of two different classes in class C of methods in class D, and polymorphism is considered into account</td>
</tr>
<tr>
<td>PIM [8]</td>
<td>This method calculates the number of invocations in class C of methods that invoke methods of another class</td>
</tr>
<tr>
<td>PIMAS [8]</td>
<td>This method calculates the number of direct or indirect invocations between static methods, and polymorphism is taken into account</td>
</tr>
<tr>
<td>INAG [8]</td>
<td>The metric is 1 if there is an indirect aggregation between two classes; otherwise, the metric is 0</td>
</tr>
<tr>
<td>ACAIC [15]</td>
<td>$ACAIC(c) = \sum_{d \in \text{Ancestors}(c)} \text{CA}(c, d)$; this calculates the number of out-degrees between one class and attributes of another class with inheritance</td>
</tr>
<tr>
<td>OCAIC [15]</td>
<td>$OCAIC(c) = \sum_{d \in \text{Others}(c) \cap \text{Friends}(c)} \text{CA}(c, d)$; this calculates the number of out-degrees between one class and attributes of another class without inheritance</td>
</tr>
<tr>
<td>ACMIC [15]</td>
<td>$ACMIC(c) = \sum_{d \in \text{Ancestors}(c)} \text{CA}(c, d)$; it calculates the number of out-degrees between one class and methods of another class with inheritance</td>
</tr>
<tr>
<td>OCMIC [15]</td>
<td>$OCMIC(c) = \sum_{d \in \text{Others}(c) \cap \text{Friends}(c)} \text{CA}(c, d)$; it calculates the number of out-degrees between one class and methods of another class without inheritance</td>
</tr>
<tr>
<td>AMMIC [15]</td>
<td>$AMMIC(c) = \sum_{d \in \text{Ancestors}(c)} \text{MM}(c, d)$; it calculates the number of out-degrees for methods between two classes with inheritance</td>
</tr>
<tr>
<td>OMMCIC [15]</td>
<td>$OMMICIC(c) = \sum_{d \in \text{Others}(c) \cap \text{Friends}(c)} \text{MM}(c, d)$</td>
</tr>
<tr>
<td>ICF, FCF [16]</td>
<td>$ICF_i = \sum_{k=1}^{n} I(k, i) \text{ICF}<em>j, I(i, j) = (i, j) \sum</em>{h=1}^{n} e(i, k), FCF_j = \sum_{k=1}^{n} F(k, i) \text{FCF}<em>j, I(i, j) = (i, j) \sum</em>{h=1}^{n} e(k, j)$</td>
</tr>
</tbody>
</table>

Direction. Aggregation implies that one class exists in another class in the form of attribute.

**Definition 2.** DEP relationship (dependency)

DEP_D: dynamic dependency refers to an instance method in a class that invokes methods and attributes in another class.

DEP_S: static dependency refers to static methods in a class invokes methods and attributes in another class.

**Definition 3.** GEN (generalization)

One class inherits with another class, or one class implements interfaces with another class, or that of an abstract class.

2.2.2. Definition of Package Hierarchy. Packages of an object-oriented software system include classes and subpackages in the current hierarchy, and these subpackages contain classes in the current hierarchy and their subpackages. Software systems can actually be considered to be a tree hierarchical structure composed of packages.

Layer $l$ of a package is defined as $p^l \leq \mathbb{E}^{l+1}, R^{l+1} > . \mathbb{E}^{l+1}$ represents a set of classes in the layer, while this layer does not contain subpackages of this layer. $R^{l+1}$ represents class relationship in the layer, that is, $R^{l+1} \subseteq \mathbb{E}^{l+1} \times \mathbb{E}^{l+1}$.

Layer $l$ of a package is defined as a set of weighted fan-out of $C_i$ at the layer of the package, that is, $S_{layer}(i) \subseteq R^{l+1}$. $S_{all}(i)$ is the set of weighted cross-package fan-out of $C_i$, that is, $S_{all}(i) \subseteq R^{l+1} \cup R^{l+2} \cup R^{l+3} \cdots$.

2.3. CSBG for Coupling Measurement. Software stability and modularity could be measured based on complex network theories. In this study, software systems can be expressed as a set of bipartite graphs that use nodes as classes, and ASS as well as DEP are the edges constituted by attributes of the class with those of another class based on complex network theories. However, GEN is a direct connection between
<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Strength</th>
<th>Fan-out/ fan-in</th>
<th>Indirect coupling</th>
<th>Indirect Inheritance</th>
<th>Weighted Static invocation</th>
<th>System metric</th>
<th>Package level</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBO</td>
<td>Method invocation, attribute reference</td>
<td>No</td>
<td>No</td>
<td>Both</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>CBO'</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RFC</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RFC'</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>MPC</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DAC</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>COP</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>COF</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>IH-ICP</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>NH-ICP</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SIMAS</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>PIM</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>PIMAS</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>INAG</td>
<td>Type of attribute</td>
<td>Yes</td>
<td>Both</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ACAIC</td>
<td>Type of attribute</td>
<td>Yes</td>
<td>Both</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>AMMC</td>
<td>Method invocation</td>
<td>Yes</td>
<td>Both</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>AMCIC</td>
<td>Method invocation</td>
<td>Yes</td>
<td>Both</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>OCF</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>OCMIC</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SIMAS</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>AMMC</td>
<td>Method invocation</td>
<td>Yes</td>
<td>Both</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>OCF</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>OCMIC</td>
<td>Method invocation</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 3: Comparative analysis for the typical methods of coupling measurement.
classes. Therefore, object-oriented software systems are taken into account as a set of special bipartite graphs constituted by classes in the package, as shown in Figure 1. There are defects in the coupling metrics containing the two metrics of fan-in and fan-out, because their total number is equal in a software system [46]. Therefore, this study only analyzed fan-out metric. The coupling strength between classes is correlated with the complexity of information exchange between modules. The more complex the information interaction (such as CBO), the tighter the coupling [47]. Coupling measurement metrics refer to the weighted fan-out of classes in special bipartite graphs. In these special bipartite graphs, classes associate with a class that may be at the same layer of the same package or at different layers of the package. Therefore, this study analyzed degrees of fan-out for classes in the same layer and all layers of the package. Moreover, heterogeneity of the abovementioned weighted out-degree was analyzed.

2.3.1. Demonstration of CSBG for Coupling Measurement. The detailed scheme proposed here is explained in the following, as illustrated in Figure 2:

(1) The object-oriented software systems are constructed as directed weighted network graphs, and classes and relationship between classes are shown as nodes and edges, respectively.

(2) The package level, class level, and method level are combined to construct special weighted bipartite graphs between classes, aiming to make preparation for calculating the weighted out-degree of classes at the same layer of the package (see step 3, \((\sum S_{\text{layer}}(i)))\), and the weighted out-degree \((\sum S_{\text{all}}(i)))\) of all classes with classes across layers of the package (see step 3).

(3) \(S_{\text{layer}}\), \(DEP_{\text{Dlayer}}\), \(DEP_{\text{Slayer}}\), and \(GEN_{\text{layer}}\) at the same layer of the package were calculated. \(S_{\text{layer}}\) is determined by adding the weights of \(x_1, x_2, x_3, x_4\) to the four mentioned metrics, respectively, in order to calculate \(ASS_{\text{all}}\), \(DEP_{\text{Dall}}\), \(DEP_{\text{all}}\), and \(GEN_{\text{all}}\) between classes across different layers of the package. Then, weights of \(x_5, x_6, x_7, x_8\) were added to these four metrics to determine \(S_{\text{all}}\):

\[
\begin{align*}
\sum_{i=1}^{n} S_{\text{layer}}(i) &= \left[ \sum_{i=1}^{n} ASS_{\text{layer}}(i) \right] \left[ \sum_{i=1}^{n} DEP_{\text{Dlayer}}(i) \right] \\
&\quad \times \left[ x_1, x_2, x_3, x_4 \right]^T, \\
\sum_{i=1}^{n} S_{\text{all}}(i) &= \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} DEP_{\text{Dall}}(i) \right] \\
&\quad \times \left[ x_5, x_6, x_7, x_8 \right]^T.
\end{align*}
\]

(4) \(S_{\text{layer}}\) and \(S_{\text{all}}\) are weighted to calculate the weighted out-degree \(S\) of the software system. The system coupling is calculated through dividing \(S\) by the number of classes:

\[
S = \frac{\sum_{i=1}^{n} S(i)}{n} = \left[ \sum_{i=1}^{n} S_{\text{layer}}(i), \sum_{i=1}^{n} S_{\text{all}}(i) \right] \times [\alpha, \beta]^T,
\]

\[
\overline{S} = \frac{S}{n}
\]

2.3.2. Calculation of the Weighted Fan-Out of Classes in a Software System. A special bipartite graph is constructed between classes of a software system. Weighted fan-out of classes in the special bipartite graph is analyzed based on characteristics of the object-oriented software.

(1) Construction of the Special Bipartite Graph in Software Systems. In the graph \(G(V, E)\), if we divide the set \(V\) of nodes into two complementary subsets \(S\) and \(T\), \(S \cup T = V\), and \(S \cap T = \phi\), the graph \(G(V, E)\) is the bipartite graph. In the graph \(G_{ij}(C_{pi}, E_{ij})\) constructed by classes \(C_i\) and \(C_j\) for the software, if only coupling relationship between classes is considered, coupling of methods and attributes in the class wouldn’t be taken into account; then, the
property of bipartite graph $C_i \cap C_j = \emptyset$ is satisfied. A network diagram constructed by classes $C_i$ and $C_j$ satisfies the following formula: $C_i \cup C_j = C_{ij}$. Moreover, the two points of a connecting edge between classes $C_i$ and $C_j$ are in classes $C_i$ and $C_j$, respectively.

In summary, the complex coupling between classes $C_i$ and $C_j$ constructs a bipartite graph $G_{ij}(C_{ij}, E_{ij})$. However, $G_{ij}(C_{ij}, E_{ij})$ not only possesses the general properties of a bipartite graph, including method invocation and dependencies, but also possesses its own characteristics. In aggregation, reference, inheritance, and interface implementation between classes $C_i$ and $C_j$, the two points of the connecting edge are in classes $C_i$ and $C_j$, respectively. This bipartite graph $G_{ij}(C_{ij}, E_{ij})$ is defined as a special bipartite graph. However, the software system $G(C, E)$ can be considered as a set of special bipartite graphs $G_{ij}(C_{ij}, E_{ij})$ as well (Figure 3).

In the present study, the coupling of the complete bipartite graph $G_{ij}(C_{ij}, E_{ij})$ constructed by classes $C_i$ and $C_j$ was used to analyze the coupling of the software system $G(C, E)$.

(2) Modeling the Coupling of Special Bipartite and Calculating the Number of Weighted Fan-Out in Software Systems. In this study, a software system $G = (C_1, C_2, \ldots, C_N)$ was defined. Classes $C_i$ and $C_j$ were defined as two different classes in a software system: $C_j = (O_j, A_j, M_j)$. Among them, $O_j = \{O_{j1}, O_{j2}, \ldots, O_{jp}\}$ was the set of instantiated objects in class $C_j$ and $p$ is the number of instantiated objects. $A_j = \{A_{j1}, A_{j2}, \ldots, A_{jq}\}$ is the attribute set of class $C_j$, and $q$ is the number of attribute. $M_j = \{M_{j1}, M_{j2}, \ldots, M_{jr}\}$ is the method set of class $C_j$, and $r$ is the number of methods. The methods included class methods and instance methods, that is, $(C_j \setminus C_j \cap M_j) \subseteq M$.

The relationship of the special bipartite graph between classes $C_i$ and $C_j$ can be summarized as follows:

**ASS**

In class $C_i$, there was instantiation of class $C_j$ (association), or one class existed in another class in the form of attribute (aggregation), which was defined as $C_j \cdot O_{jp}$, where $1 \leq p \leq p$.

In the class $C_i$, instantiated object $O_{jp}$ of class $C_i$ was implemented ($1 \leq p \leq p$), or $O_{jp}$ was a part of class $C_j$; then there was an ASS edge between classes $C_i$ and $C_j$ that is, $(C_i, C_j, O_{jp}) \in R_{ASS}$. The set of ASS weighted fan-out of class $C_i$ was

$$\text{ASS}(i) = \left\{ (C_i, C_j, O_{jp}) \mid 1 \leq j \leq N, 1 \leq p \leq p, (C_i, C_j, O_{jp}) \in R_{ASS} \right\}.$$  

**DEP**

Relationship between classes is implemented defined by instance methods and variables. In class $C_i$, there are instance methods of class $C_j$: if $C_j \cdot O_{jp} \cdot M_{jr}$, $1 \leq j \leq N, 1 \leq p \leq p$, and $1 \leq r \leq r$, then the relationship between classes $C_i$ and $C_j$ is defined as $(C_i, C_j, O_{jp} \cdot M_{jr}) \in R_{DEP_D/DEP_S}$. In class $C_j$, there are instance variables of class $C_j$: if $C_j \cdot O_{jp} \cdot A_{jq}$, $1 \leq j \leq N, 1 \leq p \leq p$, and $1 \leq q \leq q$, then relationship between classes $C_i$ and $C_j$ is defined as $(C_i, C_j, O_{jp} \cdot A_{jq}) \in R_{DEP_D/DEP_S}$. Under the condition of instance methods and instance variables, the set of weighted fan-out for $C_i$ was

\[ \sum_{S(i)} \sum_{\text{ASS}(i)} \sum_{\text{DEP}_D(i)} \sum_{\text{DEP}_S(i)} \sum_{\text{GEN}(i)} \]
Implementing connecting edges between two classes through class methods and class variables. If there are class methods of \( C_j \) (static methods) in class \( C_i \) and \( 1 \leq r' \leq r \), then the relationship between classes \( C_i \) and \( C_j \) is defined as \((C_i, C_j, M_{j'}', A_{j'q}) \in R_{DEP.S-A,M} \). When there are class variables (static variables) of \( C_j \) in class \( C_i \), then the relationship between classes \( C_i \) and \( C_j \) is defined as \((C_i, C_j, A_{j'q}) \in R_{DEP.S-A} \). Thereafter, under the conditions of class methods and class variables, the set of weighted out-degree for class \( C_i \) was

\[
DEP.D(i) = \{ (C_i, C_j, O_{j'p}, M_{j'}), (C_i, C_j, O_{j'p}, A_{j'q}) \} \\
\times (1 \leq j \leq N, 1 \leq p \leq p_i, 1 \leq r' \leq r, \\
1 \leq q' \leq q_i, (C_i, C_j, O_{j'p}, M_{j'}) \in R_{DEP.D}, \\
(C_i, C_j, O_{j'p}, A_{j'q}) \in R_{DEP.D-A} \}. 
\]

(4)

**GEN**

As the inheritance is preferred in software engineering, if one class is a subclass of another class, the derived connecting edge was taken into account only once in this study. Because transfer of derived connecting edges would make the software system network more complex, this study did not consider transfer of derived connecting edges but only considered the conditions that class \( C_i \) was a direct subclass of class \( C_j \) (through extension), or class \( C_i \) was implemented through interface class \( C_j \) (through implementation), or class \( C_i \) was implemented by abstract class \( C_j \) (through extension). Thus, there was a GEN connecting edge between classes \( C_i \) and \( C_j \), which was defined as \((C_i, C_j) \in R_{GEN} \), and a set of GEN weighted out-degree for class \( C_i \) was

\[GEN(i) = \{ (C_i, C_j) \mid (C_i, C_j) \in R_{GEN} \}.\]

(6)

(3) **Determination of Weights.** In software systems, one class can construct one or more special bipartite graphs with other classes. Supposing that the number of classes and the total number of weighted fan-out of all classes are definite in a software system, the first case is that the number of weighted fan-out in each class is the same or roughly the same. The second case is that there is no rule for the distribution of the number of weighted fan-out in a class. The third case is that the number of weighted fan-out of a class is heterogeneity, which approximately obeys the power-law distribution. For the second case, heterogeneity of the out-degree of the class is superior than that of the first case; however, this is impossible to be compared with the third case. For the third case, because the number of fan-out is limited for the majority of classes, only few classes have a large number of fan-out; therefore, maintenance staff can dedicate more effort on these few classes. Moreover, the maintenance workload of these classes is lower than that of the first case.

In this study, heterogeneity under the situation of fan-out was analyzed. If the distribution was the above-mentioned third case, then the larger the power-law value, the easier the maintenance, and the smaller the coupling degree. However, if the distribution was one of the other two cases, then it was considered in this study that the power-law value was equal to 1. \( b_1, b_2, b_3, \) and \( b_4 \) are the power-law values for the distribution of \( ASS \)-layer, \( DEP.D \)-layer, \( DEP.S \)-layer, and \( GEN \)-layer, respectively. \( b_5, b_6, b_7, \) and \( b_8 \) are the power-law values for the distribution of \( ASS \)-all, \( DEP.D \)-all, \( DEP.S \)-all, and \( GEN \)-all, respectively. The calculating formula for weights was as follows:

\[
[x_1, x_2, x_3, x_4]^T = \frac{4 b_1 + 4 b_1 + b_2 + b_2 + b_2 + b_2 + b_2 + b_2}{b_1 + b_1 + b_2 + b_2 + b_2 + b_2 + b_2 + b_2},
\]

\[
[x_5, x_6, x_7, x_8]^T = \frac{8 b_5 + 8 b_5 + 8 b_6 + 8 b_6 + 8 b_6 + 8 b_6 + 8 b_6 + 8 b_6}{b_5 + b_5 + b_6 + b_6 + b_6 + b_6 + b_6 + b_6},
\]

\[
\alpha = \frac{\sum_{i=1}^{8} b_i}{\sum_{i=1}^{8} b_i},
\]

\[
\beta = 1 - \alpha.
\]

In the present study, statistical analyses were performed for the out-degree of the three open-source software systems, and the distributions were the first and the third cases as mentioned above, demonstrating that the proposed method had a certain practical value.

### 2.4. Theoretical Verification of Coupling Metrics

Whether the proposed CSBG method met the mathematical properties of the measurement metrics [4] was theoretically verified.

**CSBG Property 1.** CSBG satisfies nonnegativity.

**Proof.** In an object-oriented software system \( G = (\{c_1, c_2, \ldots, c_N\}) \), there are two classes \( c_1, c_2 \in G \). When \( ASS \)-layer, \( DEP.D \)-layer, \( DEP.S \)-layer, and \( GEN \)-layer are all 0, the minimum value of the software system CSBG \((G)\) is 0. However, there is a maximum value \( M (M > 0) \), so that the CSBG \((G)\) value is in the range of \([0, M]\). Thus, nonnegativity of CSBG is satisfied, and the proposition is proved.\[\square\]

**CSBG Property 2.** CSBG satisfies zero value.
Proof. As described in CSBG property 1, if the minimum value is 0, then CSBG satisfies zero-value, and the proposition is proved as well.

CSBG Property 3. CSBG satisfies monotonicity.

Proof. If one edge is arbitrarily added in the system, the weighted out-degree of classes would increase according to CSBG measurement metrics. Obviously, the coupling increases as well. Thus, CSBG meets monotonicity and the proposition is proved.

CSBG Property 4. CSBG meets the property of class merging.

Proof. In an object-oriented software system $G = (c_1, c_2, \ldots, c_N)$, there are two classes $c_1, c_2 \in G$, and class $c'$ is a merger of classes $c_1$ and $c_2$. The object-oriented system $G'$ is a system in which classes $c_1$ and $c_2$ in $G$ are replaced by class $c'$. CSBG mainly calculates the weighted out-degree of classes in software systems. Therefore,

$$[\text{CSBG}(c_1) + \text{CSBG}(c_2) \geq \text{CSBG}(c') \mid \text{CSBG}(G) \geq \text{CSBG}(G')]$$

(8)

CSBG Property 5. CSBG satisfies the merge property of two irrelevant classes.

Proof. In an object-oriented software system $G = (c_1, c_2, \ldots, c_N)$, there are two classes $c_1, c_2 \in G$, and the two classes are not coupled. Moreover, class $c'$ is the merger of classes $c_1$ and $c_2$. The object-oriented system $G'$ is a system in which classes $c_1$ and $c_2$ in $G$ are replaced by class $c'$. CSBG mainly calculates the weighted out-degree of classes in software systems. Therefore,

$$[\text{CSBG}(c_1) + \text{CSBG}(c_2) = \text{CSBG}(c') \mid \text{CSBG}(G) = \text{CSBG}(G')]$$

(9)

2.5. Comparative Experiment. In the next sections, the CSBG method is herein proposed for coupling measurement and the existing measurement methods were compared and analyzed in order to verify the rationality of the results of CSBG measurement.

2.5.1. Calculating the Coupling of the Software System Using CSBG. In this section, CSBG for coupling measurement was compared with the existing measurement methods.

This experiment was conducted on a simple system as an example to analyze and compare the measurement values by the existing coupling measurements. This system was composed of 6 classes (Shape.java, Point.java, Line.java, Triangle.java, Quadrilateral.java, and Square.java), which described shapes, points, edges, triangles, quadrilaterals, and squares, respectively.

Among them, the first three classes were in package graph, and the last three classes were in package graph.polygon (hierarchy of classes in package level is shown in Figure 4). There were inheritance, combination, variable declaration, and method invocation among these classes, which were appropriate for analyzing the coupling model. Codes of classes are shown in Figures 5–10.
There were three classes in the package graph, including class Shape, class Point, and class Line.

There were three classes in the package graph.polygon, which were classes of Triangle, Square, and Quadrilateral.

In this study, an algorithm was designed and the program was developed based on the aforementioned mathematical model, mainly calculating the four metrics for the out-degree of classes in the same layer and different layers of the package in software systems, including ASS, DEP_D, DEP_S, and GEN. Coupling metrics, including ASS, DEP_D, DEP_S, and GEN, were corresponded to the cases described in Section 2.3. Out-degrees of classes in various layers are shown in Table 4.

The mathematical model described in Section 2.3 was herein used. Because the number of classes was small, the heterogeneity of out-degree of classes could not be reflected. Moreover, heterogeneity had little impact on the coupling in this example. Therefore, it was considered that heterogeneity was approximately the same. Coupling of software systems was calculated as follows:
Figure 10: Demonstration of class quadrilateral.

\[ CSBG(G) = \frac{\sum_{i=1}^{6} CSBG(C_i)}{6} = 6.55, \]

\[ [x_1, x_2, x_3, x_4]^T = [1, 1, 1, 1]^T, \]

\[ [x_5, x_6, x_7, x_8]^T = [1, 1, 1, 1]^T, \]

\[ \alpha = 0.5, \]

\[ \beta = 1 - \alpha = 0.5, \]

\[ \sum_{i=1}^{n} S_{\text{layer}}(i) = \left[ \sum_{i=1}^{n} S_{\text{all}}(i) \right] \times [x_1, x_2, x_3, x_4]^T \]

\[ = [6, 10, 0, 3] \times [1, 1, 1, 1]^T \]

\[ = 19, \]

\[ \sum_{i=1}^{n} S_{\text{all}}(i) = \left[ \sum_{i=1}^{n} S_{\text{layer}}(i) \right] \times [x_5, x_6, x_7, x_8]^T \]

\[ = [24, 38, 0, 5] \times [1, 1, 1, 1]^T \]

\[ = 67, \]

\[ S = \sum_{i=1}^{n} S(i) = \left[ \sum_{i=1}^{n} S_{\text{layer}}(i) \right] \times [\alpha, \beta]^T \]

\[ = [19, 67] \times [0.5, 0.5]^T \]

\[ = 43, \]

\[ \bar{S} = \frac{S}{n} = \frac{43}{6} = 7.17. \]
measurement contained a certain rationality in theory. Therefore, the CSBG classes and the special bipartite graph composed of classes from object-oriented software systems but also analyze the complexity of coupled software systems from an overall and global perspective. Therefore, the measurement values were mostly either very large or very small, and several metrics were equal to 0. In addition, discrimination of these metrics was not significant compared with other methods for coupling measurement. The metrics calculated by CSBG had a better discrimination. Therefore, the existed methods have limitations, which cannot accordingly satisfy an effective coupling measurement for complex software systems. CSBG not only can consider a complex relationship between classes in object-oriented software systems but also analyze the complexity of classes and the special bipartite graph composed of classes from the perspective of overall package level. Therefore, the CSBG measurement method contained a certain rationality in theory.

### 3. Results


In order to further validate the effects of CSBG, this study used CSBG to measure and analyze coupling between classes in the three Java open-source software systems from different fields, including Art of Illusion [48], JabRef [49], and GanttProject [50]. Some studies have reported results of class cohesion metrics for the three open-source software systems [51–53]; it is feasible to further study the complexity of the three open-source software systems if there is a more reasonable method for coupling measurement. In order to verify the effects of the CSBG measurement method in actual open-source software systems, three Java open-source software systems from different aforementioned fields were used. Art of Illusion is a software system for 3D rendering, modeling, and animation. JabRef is a graphical application for managing bibliographic database. GanttProject is a software system for project scheduling characterized by resource calendar, management, and import/export (MS Project, PDF, HTML). The reasons to use the three open-source software systems in the measurement were because (1) these systems were based on object-oriented Java; (2) the classes in the systems had a certain scale; (3) the three systems were from different fields; and (4) the source codes were available as well. Scholars can freely download the source codes from an open-source website (http://sourceforge.net).


Firstly, the program was developed and out-degree of classes at the same layer and all layers of the package was eventually obtained, including ASS, DEP, DEP, and GEN.

In this section, DEP and DEP were analyzed, and the results are shown in Figures 11–18. In the experimental results, DEP was a nonstandardized part of probability distribution $P(i)$. If, $P(i) \sim i^γ$, then $DEP \sim (i)^γ$. A linear function was fitted using the double logarithmic method that was fitted to estimate Gamma index $γ$ ($R$ is the Pearson’s correlation coefficient and $SD$ is standard deviation; $γ$ is also expressed as $B$ in the following table).

Although inheritance between classes increases coupling of the system, this is encouraged by the software system, which is conducive to reduce function definition and attribute definition in order to create a new class; thus, it is a poor coupling. It can be seen from linear distribution of GEN (Table 6) that neither all classes have an inheritance relationship, nor the GEN fan-out of all classes were very large or very small. However, classes with values equal to 0 or 1 were dominant.

Pearson’s correlation coefficient ($R$) and SD value provided the quality of the linear fitting; the larger the R value, the better the quality of the linear fitting, and $B$ is estimated Gamma index $γ$. Moreover, the smaller the SD value, the better the quality of the linear fitting. As shown in Table 7, if 0.95 is considered to be the minimum value, it can be approximated that the distribution obeyed the power-law distribution except that ASS value in JabRef was relatively small (0.91651 and 0.88148). Furthermore, the distributions of ASS, ASS, DEP, DEP, DEP, DEP, and DEP all layer were assumed to obey power-law distribution. The results demonstrated that there was a certain rule for the number of fan-out of classes in the form of ASS and DEP, which was not the case that the values were mostly large or small. However, they had “scale-free” property for complex networks, which obeyed the power-law distribution. In actual software development process, if software developers excessively pursue low coupling between classes, a class may be divided into two or more subclasses; thus, system complexity may be accordingly increased. The process of determination of the range of coupling

<table>
<thead>
<tr>
<th>Class name</th>
<th>Out-degree at the same layer</th>
<th>Out-degree of all layers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ASS</td>
<td>DEP $D$</td>
</tr>
<tr>
<td>Quadrilateral</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Triangle</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Point</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shape</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Square</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>10</td>
</tr>
</tbody>
</table>
Table 5: Results of various methods for coupling measurement.

<table>
<thead>
<tr>
<th>Quadrilateral</th>
<th>Triangle</th>
<th>Line</th>
<th>Point</th>
<th>Shape</th>
<th>Square</th>
<th>Software system</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSBG</td>
<td>17.5</td>
<td>13.3</td>
<td>5.1</td>
<td>0.3</td>
<td>0</td>
<td>3.1</td>
</tr>
<tr>
<td>CBO</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>CBO'</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RFC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RFC'</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MPC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DAC</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>DAC'</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>COF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICP</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IH-ICP</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NIH-ICP</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SIMAS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PIM</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PIMAS</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>INAG</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ACAIC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OCAIC</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ACMIC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OCMIC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AMMC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OMMC</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ICF</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FCF</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

(a) \(y = 2.51614 \times 10 - 5.48165\) 
\(R = 0.97722\) 
\(SD = 0.08465\) 
\(1 < i < 471\)

(b) \(y = 7.1483 \times x - 19.22832\) 
\(R = 0.91651\) 
\(SD = 0.17962\) 
\(0 < i < 695\)

(c) \(y = 7.22368 \times x - 20.54354\) 
\(R = 0.94567\) 
\(SD = 0.11961\) 
\(0 < i < 947\)

Figure 11: The double logarithmic diagrams of the fan-out of ASS invocation for classes at the same layer of a package.

(a) \(y = 6.07342 \times x - 14.40843\) 
\(R = 0.97249\) 
\(SD = 0.13761\) 
\(0 < i < 471\)

(b) \(y = 20.06268 \times x - 55.32035\) 
\(R = 0.97537\) 
\(SD = 0.12677\) 
\(0 < i < 695\)

(c) \(y = 25.12239 \times x - 73.32253\) 
\(R = 0.95689\) 
\(SD = 0.14931\) 
\(0 < i < 947\)

Figure 12: The double logarithmic diagrams of the fan-out of DEP_D invocation for classes at the same layer of a package.
Figure 13: The double logarithmic diagrams of the fan-out of ASS invocation for classes at all layers of a package.

Figure 14: The double logarithmic diagrams of fan-out of GEN invocation for classes at the same layer of a package.

Figure 15: The double logarithmic diagrams of the fan-out of ASS invocation for classes at all layers of a package.

Figure 16: The double logarithmic graph of fan-out of DEP_D invocation for classes at all layers of a package.
Figure 17: The double logarithmic diagrams of fan-out of DEP_S invocation for classes at all layers of a package.

Figure 18: The double logarithmic diagrams of fan-out of DEP_S invocation for classes at all layers of a package.

Table 6: Values of fan-out for different classes of GEN.

<table>
<thead>
<tr>
<th>Value of fan-out</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEN_layer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>244</td>
<td>199</td>
<td>21</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>JabRef</td>
<td>479</td>
<td>192</td>
<td>16</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GanttProject</td>
<td>630</td>
<td>264</td>
<td>45</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GEN_all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>173</td>
<td>246</td>
<td>29</td>
<td>19</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>JabRef</td>
<td>309</td>
<td>276</td>
<td>79</td>
<td>27</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GanttProject</td>
<td>485</td>
<td>349</td>
<td>79</td>
<td>21</td>
<td>10</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7: Values of R, SD, and B parameters.

<table>
<thead>
<tr>
<th>Software system</th>
<th>R</th>
<th>SD</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASS_layer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>0.9772</td>
<td>0.08465</td>
<td>2.51614</td>
</tr>
<tr>
<td>JabRef</td>
<td>0.91651</td>
<td>0.17962</td>
<td>7.1483</td>
</tr>
<tr>
<td>GanttProject</td>
<td>0.94567</td>
<td>0.11961</td>
<td>7.22368</td>
</tr>
<tr>
<td>DEP_D_layer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>0.97249</td>
<td>0.31761</td>
<td>6.07342</td>
</tr>
<tr>
<td>JabRef</td>
<td>0.97537</td>
<td>0.12677</td>
<td>20.06268</td>
</tr>
<tr>
<td>GanttProject</td>
<td>0.95689</td>
<td>0.14931</td>
<td>25.12239</td>
</tr>
<tr>
<td>DEP_S_layer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>0.95388</td>
<td>0.14533</td>
<td>5.52632</td>
</tr>
<tr>
<td>JabRef</td>
<td>0.94658</td>
<td>0.19504</td>
<td>14.97997</td>
</tr>
<tr>
<td>GanttProject</td>
<td>0.94086</td>
<td>0.1309</td>
<td>17.20075</td>
</tr>
<tr>
<td>ASS_all</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>0.97414</td>
<td>0.10346</td>
<td>1.84955</td>
</tr>
<tr>
<td>JabRef</td>
<td>0.88148</td>
<td>0.2079</td>
<td>1.65033</td>
</tr>
<tr>
<td>GanttProject</td>
<td>0.94905</td>
<td>0.13485</td>
<td>3.64204</td>
</tr>
<tr>
<td>DEP_D_all</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>0.97028</td>
<td>0.14556</td>
<td>3.39056</td>
</tr>
<tr>
<td>JabRef</td>
<td>0.97314</td>
<td>0.13056</td>
<td>8.47787</td>
</tr>
<tr>
<td>GanttProject</td>
<td>0.96602</td>
<td>0.13634</td>
<td>9.56684</td>
</tr>
<tr>
<td>DEP_S_all</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illusion</td>
<td>0.97265</td>
<td>0.1208</td>
<td>3.5066</td>
</tr>
<tr>
<td>JabRef</td>
<td>0.96831</td>
<td>0.14207</td>
<td>4.47638</td>
</tr>
<tr>
<td>GanttProject</td>
<td>0.95419</td>
<td>0.12876</td>
<td>5.99337</td>
</tr>
</tbody>
</table>
between classes in software systems is significant. Based on data analysis, it can be seen that scale-free property of complex networks motivated software developers to pay more attention to the distribution range of the coupling in large-scale software systems, which could provide a reliable reference for developing more reasonable software systems.

3.3. Coupling Measurement for the Three Open-Source Software Systems. According to the results of the above-mentioned analysis, out-degrees of classes were often equal to 0, 1, and 2 for class inheritance in generalization, interface implementation, and implementation of abstract classes, which were approximately linearly distributed. Therefore, the power-law value of GEN was approximated to 1.

3.3.1. Calculation of Coupling Measurement for Art of Illusion. According to the CSBG method for coupling measurement, coupling of the software system for Art of Illusion was calculated as follows:

\[ [x_1, x_2, x_3, x_4]^T = \left[ \sum_{i=1}^{4} b_{1} \sum_{i=1}^{4} b_{2} \sum_{i=1}^{4} b_{3} \sum_{i=1}^{4} b_{4} \right]^T \]

\[ = \left[ \sum_{i=1}^{4} 2.51614 \sum_{i=1}^{4} 6.07342 \sum_{i=1}^{4} 5.52632 \sum_{i=1}^{4} 1 \right]^T \]

\[ = [6.007567, 2.48858, 2.735252, 15.11588]^T, \]

\[ [x_5, x_6, x_7, x_8]^T = \left[ \sum_{i=5}^{8} b_{1} \sum_{i=5}^{8} b_{2} \sum_{i=5}^{8} b_{3} \sum_{i=5}^{8} b_{4} \right]^T \]

\[ = \left[ \sum_{i=5}^{8} 1.84955 \sum_{i=5}^{8} 3.39056 \sum_{i=5}^{8} 3.5066 \sum_{i=5}^{8} 1 \right]^T \]

\[ = [5.269774, 2.874661, 2.779533, 9.74671]^T, \]

\[ \alpha = \frac{\sum_{i=1}^{n} b_i}{\sum_{i=1}^{n} b_i} = 0.39023, \]

\[ \beta = 1 - \alpha = 0.607977, \]

\[ \left[ \sum_{i=1}^{n} S_{layer}(i) \right] = \left[ \sum_{i=1}^{n} S_{ASS, layer}(i) \right] \left[ \sum_{i=1}^{n} S_{DEP_D, layer}(i) \right] \left[ \sum_{i=1}^{n} S_{GEN, layer}(i) \right] \left[ \sum_{i=1}^{n} S_{GEN, layer}(i) \right] \times [x_1, x_2, x_3, x_4]^T \]

\[ = [2479, 5848, 2005, 260] \times [6.007567, 2.48858, 2.735252, 15.11588]^T \]

\[ = 38861.91, \]

\[ \left[ \sum_{i=1}^{n} S_{all}(i) \right] = \left[ \sum_{i=1}^{n} S_{ASS, all}(i) \right] \left[ \sum_{i=1}^{n} S_{DEP_D, all}(i) \right] \left[ \sum_{i=1}^{n} S_{DEP_S, all}(i) \right] \left[ \sum_{i=1}^{n} S_{GEN, all}(i) \right] \times [x_5, x_6, x_7, x_8]^T \]

\[ = [6705, 13883, 6255, 373] \times [5.269774, 2.874661, 2.779533, 9.74671]^T \]

\[ = 96264.25, \]

\[ S = \sum_{i=1}^{n} S(i) = \left[ \sum_{i=1}^{n} S_{layer}(i) \right] \left[ \sum_{i=1}^{n} S_{all}(i) \right] \times [\alpha, \beta]^T \]

\[ = [38861.91, 96264.25] \times [0.39023, 0.607977]^T \]

\[ = 73761.21, \]

\[ S = \frac{S}{n} = \frac{73761.21}{470} = 156.9387. \]
3.3.2. Calculation of Coupling Measurement for JabRef.

According to CSBG for coupling measurement, coupling of the software system for JabRef was calculated as follows:

\[
[x_1, x_2, x_3, x_4]^T = \left[ \sum_{j=1}^4 b_1 b_j, \sum_{j=5}^4 b_1 b_j, \sum_{j=6}^4 b_1 b_j, \sum_{j=7}^4 b_1 b_j \right]^T
\]

\[
= \left[ \sum_{j=1}^4 \frac{b_j}{7.1483} \sum_{j=5}^4 \frac{b_j}{20.06268} \sum_{j=6}^4 \frac{b_j}{14.97997} \sum_{j=7}^4 \frac{b_j}{1} \right]^T
\]

\[
= [6.042129, 2.152801, 2.883247, 43.19095]^T,
\]

\[
[x_5, x_6, x_7, x_8]^T = \left[ \sum_{i=5}^8 b_i, \sum_{i=6}^8 b_i, \sum_{i=7}^8 b_i, \sum_{i=8}^8 b_i \right]^T
\]

\[
= \left[ \sum_{i=5}^8 \frac{b_i}{1.65033} \sum_{i=6}^8 \frac{b_i}{8.47787} \sum_{i=7}^8 \frac{b_i}{4.47638} \sum_{i=8}^8 \frac{b_i}{1} \right]^T
\]

\[
= [9.45543, 1.840625, 3.485982, 15.60458]^T,
\]

\[
\alpha = \frac{\sum_{j=1}^4 b_j}{\sum_{j=1}^8 b_j} = 0.26540419,
\]

\[
\beta = 1 - \alpha = 0.734596,
\]

\[
\left[ \sum_{i=1}^n S_{\text{layer}}(i) \right] = \left[ \sum_{i=1}^n \text{ASS}_{\text{layer}}(i), \sum_{i=1}^n \text{DEP}_{\text{D_layer}}(i), \sum_{i=1}^n \text{DEP}_{\text{S_layer}}(i), \sum_{i=1}^n \text{GEN}_{\text{layer}}(i) \right] \times [x_1, x_2, x_3, x_4]^T
\]

\[
= [1326, 2032, 2396, 246] \times [6.042129, 2.152801, 2.883247, 43.19095]^T
\]

\[
= 29375.8,
\]

\[
\left[ \sum_{i=1}^n S_{\text{all}}(i) \right] = \left[ \sum_{i=1}^n \text{ASS}_{\text{all}}(i), \sum_{i=1}^n \text{DEP}_{\text{D_all}}(i), \sum_{i=1}^n \text{DEP}_{\text{S_all}}(i), \sum_{i=1}^n \text{GEN}_{\text{all}}(i) \right] \times [x_5, x_6, x_7, x_8]^T
\]

\[
= [3430, 4403, 6733, 527] \times [9.45543, 1.840625, 3.485982, 15.60458]^T
\]

\[
= 71947.46,
\]

\[
S = \sum_{i=1}^n S(i) = \left[ \sum_{i=1}^n S_{\text{layer}}(i), \sum_{i=1}^n S_{\text{all}}(i) \right] \times [\alpha, \beta]^T
\]

\[
= [29375.8, 71947.46] \times [0.26540419, 0.734596]^T
\]

\[
= 60648.76,
\]

\[
\frac{S}{n} = \frac{60648.76}{694} = 87.39015.
\]
3.3.3. Calculation of Coupling Measurement for GanttProject.

According to CSBG for coupling measurement, coupling of the software system for GanttProject was calculated as follows:

\[
[x_1, x_2, x_3, x_4]^T = \left[ \sum_{j=1}^{4} b_j \sum_{i=1}^{4} b_i \sum_{j=1}^{3} b_j \sum_{i=1}^{3} b_i \right]^T
\]

\[
= \left[ 4 \cdot 7.22368 \cdot 25.12239 \cdot 17.20075 \cdot 1 \right]^T
\]

\[= [6.997378, 2.012023, 2.93864, 50.54682]^T,
\]

\[
[x_5, x_6, x_7, x_8]^T = \left[ \sum_{j=1}^{8} b_j \sum_{i=1}^{8} b_i \sum_{j=1}^{6} b_j \sum_{i=1}^{6} b_i \right]^T
\]

\[
= \left[ 8 \cdot 3.64204 \cdot 9.56684 \cdot 5.99337 \cdot 1 \right]^T
\]

\[= [5.5496, 2.111695, 3.370766, 20.20225]^T,
\]

\[
\alpha = \frac{\sum_{j=1}^{b_i}}{n} = 0.285548,
\]

\[
\beta = 1 - \alpha = 0.714452,
\]

\[
\sum_{i=1}^{n} S_{\text{layer}} (i) = \left[ \sum_{i=1}^{n} \text{ASS}_{\text{layer}} (i), \sum_{i=1}^{n} \text{DEP}_{\text{D_layer}} (i), \sum_{i=1}^{n} \text{DEP}_{\text{S_layer}} (i), \sum_{i=1}^{n} \text{GEN}_{\text{layer}} (i) \right] \times [x_1, x_2, x_3, x_4]^T
\]

\[= [1102, 1331, 565, 379] \times [6.997378, 2.012023, 2.93864, 50.54682]^T
\]

\[= 31206.69,
\]

\[
\sum_{i=1}^{n} S_{\text{all}} (i) = \left[ \sum_{i=1}^{n} \text{ASS}_{\text{all}} (i), \sum_{i=1}^{n} \text{DEP}_{\text{D_all}} (i), \sum_{i=1}^{n} \text{DEP}_{\text{S_all}} (i), \sum_{i=1}^{n} \text{GEN}_{\text{all}} (i) \right] \times [x_5, x_6, x_7, x_8]^T
\]

\[= [3201, 3833, 2291, 622] \times [5.5491, 2.111695, 3.370766, 20.20225]^T
\]

\[= 46138.17,
\]

\[
S = \sum_{i=1}^{n} S (i) = \left[ \sum_{i=1}^{n} S_{\text{layer}} (i), \sum_{i=1}^{n} S_{\text{all}} (i) \right] \times [\alpha, \beta]^T
\]

\[= [31206.69, 46138.17] \times [0.285548, 0.714452]^T
\]

\[= 41874.52,
\]

\[
\overline{S} = \frac{S}{n} = \frac{41874.52}{946} = 44.26482.
\]
The three aforementioned open-source software systems were analyzed from the points of view of package level, class level, and method level using CSBG for coupling measurement. A program was also developed to calculate various metrics; thus, the coupling of the three open-source software systems in descending order was the Art of Illusion, JabRef, and GanttProject, suggesting that it was feasible to use CSBG for coupling measurement of software systems that contained a certain practical value.

4. Conclusion

Based on bipartite graphs for complex networks, by comprehensive consideration of the weighted fan-out between classes from points of view of package level, class level, and method level, this study expressed that the interaction of classes is a special bipartite graph, while a software system is a set of these special bipartite graphs. For this purpose, first, this study analyzed the four relationships for a software system, including ASS, DEP_D, DEP_S, and GEN, and coupling relationship for a class with other classes in the same layer of package was considered as well. Moreover, coupling relationship for classes in a package with other classes in different layers of the package was taken into account. Therefore, the CSBG method for coupling measurement of software systems was proposed, which was completely in compliance with the mathematical characteristics of the widely accepted metrics. Second, for a software system, other typical methods and CSBG method were compared for the purpose of coupling measurement, and the results revealed that the measured value was either large or small due to the defects of other measurement methods that were analyzed from an overall and global perspective. Moreover, the corresponding values were mostly equal to 0. Therefore, there were some defects in other measurement methods, while CSBG had its rationality. Eventually, a program was developed based on the CSBG method to apply the three open-source software systems (Art of Illusion, JabRef, and GanttProject). The results demonstrated that coupling of the three open-source software systems in the descending order was the Art of Illusion, JabRef, and GanttProject. Although inheritance between classes increases coupling of the system, this is also followed by software engineering, which is conducive to reduce function definition and attribute definition in order to create a new class, and thus, this is weak coupling. It can be concluded from the linear distribution of GEN that all classes either had an inheritance relationship, or that the number of GEN fan-out of all classes was very large or very small. However, classes with values equal to 0 or 1 were accounted. Furthermore, it was revealed that in the same layer and total layers of the package, fan-out values of ASS, DEP_D, and DEP_S obeyed the scale-free property of complex networks. These findings provided empirical support for the CSBG method. The statistical power-law metrics were applied to the method for coupling measurement proposed in this study in order to calculate the coupling of the three open-source software systems, which provided a reliable reference for further investigation of coupling between classes in software systems using statistics of complex networks. In [54], it was mentioned that cohesion distribution of the majority classes of a software system contained a certain regularity. In other words, it was not the case that neither cohesion of all classes was very large nor very small. In the empirical analysis of coupling, the values of coupling metrics had a regularity similar to class cohesion. Although coupling represented the degree of interdependence between classes, the greater the coupling, the more complex the software from an intuitive aspect. However, excessive pursuit of “high cohesion and low coupling” of software systems increases the workload of software developers and the complexity of software systems as well. Therefore, the empirical evidence showed that within a certain range, reducing the coupling was helpful to attenuate the complexity of the software, while excessively blindly pursuit of low coupling increases the complexity of software systems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61602400). This work was also supported in part by the Key Research and Development Program of Hangzhou under Grant 20182011A46.

References


