ACADEMIC World Academic Union

ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 14, No. 4, 2019, *pp*.272-278

Ordered Metric Methods of Classes Dependency Graph Based on Structure Entropy

Aihua Gu [1 +], Jinxia Xu [1] , Lu Li [1], Shujun Li [1], Qifeng Xun [1], Jian Dong [1]and Zonglin Guo[2]

[1] *School of Information Engineering, Yancheng Teachers University, Yancheng 224002, China*
[2] *Department of Medical Instrumen,Yancheng Third People's Hospital, Yancheng 224001, China*
(Received August 01 2019, accepted September 26 2019)

Abstract. In this paper, structure entropy which is based on classes dependency graph of a software system is proposed to measure the complexity of the system. In order to research the structure entropy metrics of classes dependency graph in some large software systems, this paper takes the first step to mathematically prove that structure entropy does not have the property of cohesion. Thus, a structure entropy can be used as a single metric of complexity. And then, a program is written based on relevant matrix algorithm with the construction of classes dependency graph. The corresponding metrics of structure entropy of three pieces of open source software are calculated and figured out based on classes dependency graph. The calculation shows that most classes in the three pieces of the open source software are of randomness. Meanwhile, values of structure entropy features complex network Scale-free. Therefore, the different values of the structure entropy of open source software classes dependency graph influences the evaluation of software quality. Furthermore, some complex network statistical characteristics are found out in this paper, which will facilitate the further research on structure entropy as a metric of software complexity for sophisticated networks.

Keywords: object-oriented class, software quality, structure entropy, complex networks, software metrics

1. Introduction

A butterfly flapping its wings in tropical rain forests along Amazon River caused a tornado in an America state. This is the sentence from American meteorologist Lorentz who described the chaotic oscillator in 60s of 20th century approachably and vividly. Chaotic state can be regarded as a kind of Orderly Disorder. To combine shallow level disorder with internal regularity has called the attention of researchers. From the perspective of systematism, order is the structure form of things, that is, relationship between components of systems. Order is that whole structure of systems show regularity over time. For systematical projects, ordered structure benefits function-controllable and manageable, and facilitates system components to swap and coordinate. Therefore, Subscription services or functions can be implemented.

Software system structure has the characteristics of complex network scale-free. Heterogeneity of scale-free network represents a certain function expressing metric distribution indices of nodes. However, it may be not appropriate to use metric distribution index directly. If there are many edges in a network but very small metric distribution index, the metric of most nodes is also very large, that is, Heterogeneity becomes very little. Conversely, if there are few edges in a network but very large metric distribution index, then the metric of each node is very small. However, here heterogeneity is still very little. The inseparably link of network structure entropy and system complexity can indicate the heterogeneity degree of the network and represent the homogeneity of network nodes metric values. The heterogeneity of star network is unexpectedly the largest. Because there is only one node with very large metric in star network, in common situations, the metric values of other nodes are generally the same, so the network heterogeneity is not so strong [1].

How to represent and measure the ordered of such structure of software systems becomes the fundamental of research on complexity in science and systematic engineering relevant domains. In this paper, structure entropy is defined in terms of node degree of classes dependency graphs of software systems. Structure ordered degree of software systems is also analyzed. It provides reliable basis for optimizing overall configuration of classes in software systems.

+ Corresponding author: Aihua Gu
E-mail address: guaihua1978@163.com

A few researchers have evaluated the ordered of configuration structure and complexity basing on Shannon's information entropy defining system structure entropy, which offers solid basis for structure optimization and improvement. With the development of complex network researches, the structure entropy has been used to explore evolution rules of Internet topology structure and an order parameter which influences network topology structure evolution. These researches result that degrees of nodes in two topo structures in domain level and router level follows power law distribution [2].

In the field of software engineering, some researchers have analyzed UML graphs, and proposed entropy-based structure complexity metric method. Judging whether the system is uncertain relying on entropy can find out the system structure complexity. Zhou and his team applied entropy distance to analyze the complexity of UML graphs, when he found the relation between classes of a software system and assigned weight to edges according to the importance of classes' relation[3]. Xu and his team expressed UML graphs with weighted classes and classes dependency graphs, and judge the complexity of UML graphs according to the entropies distances[4]. Yi and his team who applied the methods proposed by Xu and Zhou verified the efficiency of the methods depending on the analysis on 27 systems[5]. Ma and his team theoretically verified the efficiency of evaluation coefficient in the metric structure order by simulating entropy decreasing process mathematically[1]. They quantitatively analyzed the order of software structures in terms of structure ordered metric computing algorism on ten different software systems and two network models. They revealed that there exists relationship between linking metric evaluation coefficient and structure entropy. With evaluation coefficient increasing, structure entropy tends decreasing. That means structure ordered degree is creasing, robustness is strengthening, systems' complexities are increasing during the period of entropy decreasing. The conclusion above offered effective proofs for further researching on different software system structure complexity. It also signifies that structure entropy metric method is feasible and effective.

Most current structure entropy analysis software systems focus on the problem of structure entropy of classes dependency graphs. This chapter analyzes the influence that structure entropy of class dependency graphs acts on software system complexity, which is directly composed of class methods and properties of three pieces of open source software.

2. Preliminaries

2.1. Structure entropy

The development of the concept of entropy of random variables and processes by Claude Shannon provided the beginning of information theory.

The entropy is the most influential concept to arise from statistical mechanics. Entropy measures the disorder in a system [6]. The theory of entropy is widely used in various fields. And the structure entropy can be used to describe the structure relation of a system.

If a system X has n sub-systems $X_1, X_2, \cdots X_n$, the sub-system X_i is related to X_j,then $g_{ij} = 1$,otherwise $g_{ij} = 0$. Let $N(i) = \sum_{j=1}^{n} g_{ij}$; linking intensity $\rho(i) = N(i)/\sum_{j=1}^{n} N(j)$ [7].

The structure entropy of system X:$H = -\sum_{i=1}^{n} \rho(i)\, ln\, \rho\,(i)$,where $\sum_{i=1}^{n} \rho(i) = 1$.

2.2. Class cohesion metric properties

According to Briand [8], class cohesion metrics carry these features: (1) non-negativity and normalization [0, Max], allowing easy comparison between various classes; (2) null and maximum values (no cohesive interactions, 0; all possible interactions within the class present, maximum); (3) monotonicity (even if the module is added cohesive interactions, the cohesion of the module cannot be decreased); (4) cohesive modules (merging two unrelated modules into one cannot enhance the cohesion of the module. Thus, for 2 classes, e.g. $c1$ and $c2$, the cohesion of the merged class c' must adhere to the following conditions: $cohesion(c') \leq max\{cohesion(c1), cohesion(c2)\}$.

3. The ordered metric methods of Classes dependency graph based on structure entropy

3.1. Dependence relations in class and dependence complex network

In the object-oriented software system, a software system is composed of a series of related classes. One class (c) describes a class in the software system, $c = (A, M)$, where $A = \{A_1, A_2, \cdots, A_a\}$is a set of attributes, and describes a attributes contained in the class c. $M = \{M_1, M_2, \cdots, M_m\}$is a set of methods, and describes m