社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

複雑ネットワークのトポロジ構造を用いたノード特性の再帰的推定手法

杉山　浩平† 　大崎　博之† 　今瀬　真†

† 大阪大学 大学院情報科学研究科
〒 565-0871 大阪府吹田市山田丘 1-5
E-mail: †{k-sugi,oosaki,imase}@ist.osaka-u.ac.jp

あらまし　本稿では、複雑ネットワークのトポロジ構造からノード特性を再帰的に推定する手法 RENC (Recursive Estimation of Node Characteristic) を提案する。これは、ネットワークのトポロジ構造からノード特性を推定するという処理を再帰的に実行することにより、ノイズの影響を低減するというものである。本稿では、さらにネットワーク生成モデル LRE (Linkage with Relative Evaluation) を提案する。ネットワーク生成モデル LRE は「人は相対評価に基づいて行動する」という性質をモデル化したものであり、ソーシャルネットワークのいくつかの特徴を再現することができる。本稿では、LRE によって生成したネットワークに RENC を適用することにより、ノード特性の再帰的推定手法 RENC の有効性を評価する。その結果、再帰的推定手法 RENC によりノード特性の推定精度が向上することを示す。
キーワード　複雑ネットワーク、トポロジ構造、ノード特性、ソーシャルネットワーク、再帰的推定手法

RENC: Recursive Estimation of Node Characteristics using Topological Structure of Complex Networks

Kouhei SUGIYAMA†, Hiroyuki OHSAKI†, and Makoto IMASE†

† Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka, 565-0871 Japan
E-mail: †{k-sugi,oosaki,imase}@ist.osaka-u.ac.jp

Abstract　In this paper, we propose a recursive estimation mechod of node characteristics called RENC (Recursive Estimation of Node Characteristics) using topological structure of complex networks. RENC reduces the effect of noise by recursively estimating node characteristics using topological structure of a network. In this paper, we also propose a network generation model called LRE (Linkage with Relative Evaluation). The network generation model LRE is for simulating a social network, in which every node is likely to make decision based on relative evaluation, so that it can reproduce several characteristics of a social network. In this paper, we evaluate the effectiveness of our recursive estimation method RENC by applying RENC to several networks generated with LRE. Consequently, we show that the estimation accuracy of node characteristics can be improved by using our recursive estimation method RENC.
Key words　Complex Networks, Topological Structure, Node Characteristic, Social Network, Recursive Estimation

1　はじめに

近年、現実に存在するさまざまな複雑ネットワークのトポロジ構造への注目が高まっている [1]。例えば、現実に存在する複雑ネットワークのトポロジ構造が、どのような規則によって決定されているのか、また、現実に存在するさまざまな複雑ネットワークが、なぜスモールワールドやスケールフリーと呼ばれる性質を持つのかについて数多くの研究がなされている [2, 3]。

また、現実の複雑ネットワークのトポロジ構造から、どのようなことが分かるかを明らかにする研究の一つとして、人と人の交流によって構築されるソーシャルネットワークに関する研究がさかんに行われている [4–6]。

さらに、ネットワーク分析手法を用いることにより、現実のさまざまな複雑ネットワークのトポロジ構造 (インターネットのトポロジ、Web のハイパーリンク構造、ブログのトラックバック関係、映画俳優の共著関係、論文の引用関係、企業の取引関係、言語における単語の共起関係など) を分析する研究もさかんに行われている [1, 7, 8]。

このような複雑ネットワークのトポロジ構造が有する可能性は、さまざまな形で応用ができると期待される。例えば、複雑

ネットワークのトポロジ構造が持つ性質を、情報検索、情報分析、データマイニングなどさまざまな用途に応用が可能であると考えられる。

例えば、従来の情報検索システムで広く用いられてきた、キーワード型の検索方式の限界が指摘されている [9]。そこで必要とする情報を検索するために、ネットワークのトポロジ構造が持つ特性をも利用した、新たな情報発見手法の登場が期待される。

複雑ネットワークのトポロジ構造から、ネットワークを構成するノードの特性 (ノード特性) を推定することができれば、これをさまざまな形で利用できる。ここでノード特性とは、ネットワークのトポロジ構造を決定する要因となるものである。例えば、企業の取引関係をネットワークとしてとらえた場合 [10]、企業の人気度・信頼度が、企業の取引関係ネットワークの構造を決定する要因となる。一般に、ソーシャルネットワークのトポロジ構造は、何らかの社会活動を反映したものである。従って、ノード (人、組織など) 特性は重要な価値を持つと考えられる。

複雑ネットワークのトポロジ構造からノード特性を推定する手法はいくつか提案されている [10–12] が、ネットワークのトポロジ構造からどのようなことが分かるか、また、ネットワークの構造をどのように活用できるかについては、これまで十分な検討が行なわれていない。

そこで本稿では、ネットワークのトポロジ構造からノード特性を再帰的に推定する手法 RENC (Recursive Estimation of Node Characteristic) を提案する。これは、ネットワークのトポロジ構造からノード特性を推定するという処理を再帰的に実行することにより、ノイズの影響を低減するというものである。

本稿では、さらにネットワーク生成モデル LRE (Linkage with Relative Evaluation) を提案する。ネットワーク生成モデル LRE は「人は相対評価に基づいて行動する」という性質をモデル化したものであり、ソーシャルネットワークのいくつかの特徴を再現することができる。本稿では、LRE によって生成したネットワークに RENC を適用することにより、ノード特性の再帰的推定手法 RENC の有効性を評価する。その結果、再帰的推定手法 RENC によりノード特性の推定精度が向上することを示す。

本稿の構成は以下の通りである。まず 2 章では、関連研究を紹介する。3 章では、ノード特性の再帰的推定手法 RENC を提案する。4 章では、ソーシャルネットワークの特徴を再現できるネットワーク生成モデル LRE を提案する。4 章では、ノード特性の再帰的推定手法 RENC の有効性を評価する。最後に、6 章において本稿のまとめと今後の課題を述べる。

2 関 連 研 究

複雑ネットワークのトポロジ構造から、ネットワークを構成するノードの特性を推定する手法に関しては、これまでいくつか研究が行われている。

文献 [11] では、Web のハイパーリンク構造をネットワークととらえ、ネットワーク構造によりノード (Web ページ) の特性を推定する手法を提案している。ここでは、ノード特性は Web ページの人気度・重要度に相当する。ネットワークのトポロジ構造から、ノード特性の指標 PageRank を推定する手法を提案している。

文献 [12] では、産業をノードとし、産業間の取引関係をリンクとした、産業間の取引関係ネットワークの構造から、ノード (産業) の特性を推定する手法を提案している。ここでは、ノード特性は産業の構造的な優位性に相当する。産業間の取引関係ネットワークにおいて、「構造的優位性」と呼ばれる指標を提案している。構造的優位性とは、産業間の取引関係ネットワークにおいて、ある産業が、その他の産業と比較して (ネットワーク構造的に) どの程度有利な位置にいるかを示すものである。

文献 [10] では、企業をノードとし、企業間の取引関係をリンクとした、企業間の取引関係ネットワークにおいて、ノード (企業) の特性を推定する手法を提案している。ここでは、ノード特性は企業の優位性に相当する。企業間の取引関係ネットワークにおいて、「構造的優位性」と呼ばれる指標を定義している。構造的優位性とは、企業間の取引関係ネットワークにおいてある企業が、その他の企業と比較して (ネットワーク構造的に) どの程度有利な位置にいるかを示すものである。

3 ノード特性の再帰型推定手法 RENC

本章では、ノード特性の再帰的推定手法である RENC (Recursive Estimation of Node Characteristic) の基本的なアイデアおよび計算アルゴリズムを説明する。さらに、構造的優位性 [10] への再帰的推定手法 RENC の適用例を説明する。

3.1 基本的なアイデアおよびアルゴリズム

2 章では、ネットワークのトポロジ構造からノード特性を推定する手法を紹介した [10–12]。これらのノード特性推定手法は、ネットワークのトポロジ構造をすべて利用してノード特性を推定している。つまり、ネットワーク中のすべてのノードと、ノード間のすべてのリンクを利用して、ノード特性を推定している。このため、ノード特性の推定において重要でない (ノード特性の推定の上ではノイズとなる) ノードやリンクを含んだままノード特性を推定していると考えられる。ネットワークのトポロジ構造をすべて利用するのではなく、ノード特性の推定にとって意味のあるノードやリンクのみを用いることにより、ノード特性の推定精度を向上させることができると考えられる。

つまり、ノード特性の再帰的推定手法のアイデアは、「ネットワークのトポロジ構造をすべて利用するのではなく、ノード特性の推定にとって意味のあるノードやリンクのみを用いてノード特性を推定する」というものである。ノード特性の推定にとって重要ではない (ノイズ) と考えられるノードやリンクを取り除いた部分ネットワークを用いて、ノード特性の推定を行う。

ここで、ノード特性の推定にとって重要でないと考えられるノードやリンクを、どのように決定するかが問題となる。このために、まず、従来と同じ方法により、ネットワークのトポロジ構造すべてを用いてノード特性の推定を行う。この結果より、ノード特性を推定する上でノイズとなると考えられるノードと、そのノードに接続されているすべてのリンクを削除する。このようにして得られた部分ネットワークに対して、再度、ノード特性の推定を再帰的に適用する。このように、ノード特性の推

定において重要でないと考えられるノードおよびリンクを順次削除しながら、再帰的にノード特性の推定を繰り返す。このような再帰的なアルゴリズムにより、ノード特性の推定におけるノイズの影響を低減することができると考えられる。

ノード特性の再帰的推定手法 RENC のアルゴリズムは以下の通りである。

(1) 従来のノード特性推定手法 (構造的優位性 [10] や PageRank [11]) を利用して、ネットワークのトポロジ構造からノード特性を推定する。

(2) 重要でないと考えられる (構造的優位性もしくは PageRank の値が小さい) 下位 T % ノードと、それらのノードに接続されているすべてのリンクをネットワークから削除する。

(3) (1) に戻る。

3.2 構造的優位性への適用例

まず、構造的優位性を簡単に説明する。

構造的優位性 SSI (Structural Superiority Index) とは、ネットワーク中のあるノードが、その他のノードと比較して (ネットワーク構造的に) どの程度優位な位置にいるかを示す指標である [10]。構造的優位性では、ネットワーク中におけるリンクが方向を持つ (つまり、有向グラフとして表現できる) 場合に、「上流のノードは下流のノードより優位である」ことを仮定し、各ノードの優位性を計算する。

具体的には、ネットワーク中のノード i の構造的優位性 C_i は次式で定義される。

$$C_i \equiv \sum_{j \to i} \zeta_{ji} (C'_j + 1) \tag{1}$$

ここで C'_j はノード j の構造的優位性の近似値であり、

$$C'_j \equiv \sum_{k \to j} \zeta_{kj} (k^I_k + 1) \tag{2}$$

と定義される。ここで、k^I_i はノード i の入次数である。また、ζ_{ij} は、ノード i からノード j にリンクが存在する場合は 1 となり、それ以外の場合は 0 となる。式 (1) は、k^I_j が大きいほど、つまり、(ノード i にリンクしている) ノード j が他の多くのノードからリンクされているほど、ノード i の構造的優位性が高くなることを意味している。

次に、ノード特性の再帰的推定手法 RENC を、構造的優位性の計算にどのように適用できるかを説明する。以下では、再帰的推定手法 RENC のアルゴリズムにおける繰り返し回数を n、n 回目の繰り返しにおいて推定された、ノード i の構造的優位性を s^n_i と表記する。

構造的優位性に RENC を適用した場合、以下のように再帰的にノードを推定する。

(0) RENC の繰り返し回数 n を $n \leftarrow 1$ と初期化する。

(1) ネットワークのトポロジ構造から構造的優位性 s^n_i を計算する。

(2) 構造的優位性 s^n_i の値が下位 T % のノードと、それらのノードに接続されているすべてのリンクをネットワークから削除する。

(3) 繰り返し回数を n を $n \leftarrow n + 1$ とし、(1) に戻る。

4 ネットワーク生成モデル LRE

本章では、ネットワーク生成モデル LRE (Linkage with Relative Evaluation) について説明する。

ネットワーク生成モデル LRE は、ソーシャルネットワークの特徴を再現するネットワークを生成するためのモデルであり、「人は相対評価に基づいて他の人と関係を構築する」という性質をモデル化している。

これまで、ネットワークを生成する、さまざまなモデルが提案されている [2, 13]。例えば、スケールフリーネットワークを生成する代表的なモデルの一つに、BA (Barabasi Albert) モデル [13] が存在する。BA モデルは、「ネットワークの成長」および「リンクの優先的選択」という性質を持つ。BA モデルでは、少数のノードにより構成されるネットワークを作成し、そこにノードを一つづつ追加してゆく。この時、新規に追加されるノードと、既存のノードとの間に、そのノードの次数に比例した確率でリンクを生成する。

また、文献 [2] では、ソーシャルネットワークの特徴を再現するネットワーク生成モデルを提案している。このモデルでも、BA モデルと同様に、ネットワークにノードを一つづつ追加するが、この時、既存のノードとの間に、そのノードの属性に比例した確率でリンクを生成する。

これらのネットワーク生成モデルにおける「リンクの優先的選択」の方法は、「新規に追加されるノードが、既存のノードを絶対評価によって判断し、リンクを生成する」と見ることができる。しかし、人間と人間の交流関係を表すソーシャルネットワークでは、絶対評価ではなく、むしろ相対評価によって関係が構築されていると考えるのが自然である。

ネットワーク生成モデル LRE では、(1) 新規に追加されたノードは、そのノードの重要度に比例した確率でリンクを生成する、(2) ノードの重要性は絶対評価ではなく相対評価によって決定される、という性質をモデル化している。これにより、ソーシャルネットワークの特徴をより再現できるネットワークが生成できると期待できる。

ネットワーク生成モデル LRE のアルゴリズムは以下の通りである。

以下では、生成したいネットワークのノード数を N、平均次数を $\overline{k}$ と表記する。また、ノード $i (1 \leq i \leq N)$ の特性値を U_i と表記する。なお、U_i は文献 [14] におけるノードの隠れ変数に相当する値である。

ネットワーク生成モデル LRE は、以下の手順によってネットワークを生成する。

(1) N 個のノードの中から、ランダムにノード i、j を選択する。

(2) 次式で定義される確率 $p_{i,j}$ により、ノード i からノード j にリンクを生成する。

$$p_{ij} = \alpha \left(\frac{U_j}{U_i + U_j} \right) \tag{3}$$

ここで α は定数である。

(3) ネットワーク中のリンク数が $\overline{k}N/2$ となればアルゴリズム

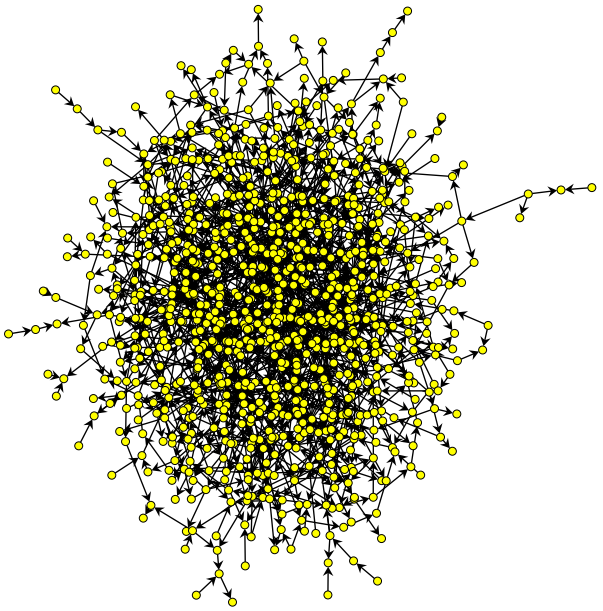

図 1: LRE によって生成したネットワークの例 ($N = 1,000, \overline{k} = 3$)

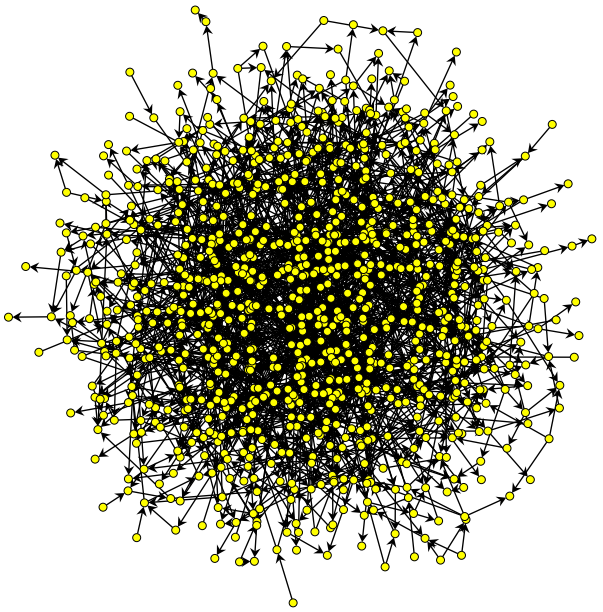

図 2: LRE によって生成したネットワークの例 ($N = 1,000, \overline{k} = 4$)

を終了する。そうでなければ (1) に戻る。

ネットワーク生成モデル LRE によって生成したネットワークの例を図 1 および図 2 に示す。図 1 は、$N = 1,000$ および $\overline{k} = 3$ とし、ノードの特性値 U_i を正規分布 $N(100, 900)$ で与えた時の結果である。また、図 2 は、$N = 1,000$ および $\overline{k} = 3$ とし、ノードの特性値 U_i を正規分布 $N(100, 900)$ で与えた時の結果である。

5 再帰的推定手法 RENC の評価

本章では、LRE によって生成したネットワークに、再帰的推定手法 RENC を適用することにより、再帰的推定手法 RENC の有効性を評価する。具体的には、再帰的推定手法 RENC により、ノード特性の推定精度がどの程度向上するかを評価する。

再帰的推定手法 RENC の各ステップにおけるノード特性の推定には、構造的優位性 (3 章) を用いた。ノード特性の推定精度を評価するための指標として、上位 k 番目までのノードの特性値 U_i の累計 S_k を用いた。

$$S_k \equiv \sum_{l=1}^{k} U_{v(l)} \tag{4}$$

ここで、$v(l)$ はノード特性値の順位が l であるノードを意味する。ノード特性の推定精度が高ければ S_k も大きな値となる。

ネットワーク生成モデル LRE を用いて、ノード数 N 1,000、平均次数 $\overline{k}$ 3 もしくは 5 のネットワークを生成した。ノードの特性値 U_i は正規分布 $N(100, 900)$ によって与えた。

平均次数 $\overline{k} = 3$ のネットワークに対して、再帰的推定手法 RENC 適用した時の結果を図 3 から図 8 に示す。また、平均次数 $\overline{k} = 5$ のネットワークに対して、再帰的推定手法 RENC 適用した時の結果を図 9 から図 14 に示す。これらの図では、再帰的推定手法 RENC の n 回目の繰り返しにおける、ノードの順位とノード特性値の累計 S_k をプロットしている。図中には、比較のため、単純に構造的優位性によって計算した場合のノードの順位とノード特性値の累計 S_k の関係もあわせて示している。

これらの図より、再帰的推定手法 RENC の繰り返し回数 n が増加するにつれ、再帰的推定手法 RENC を用いた時のノード特性値の累計 S_k が、構造的優位性を用いた時のノード特性値の累計 S_k よりも大きくなっていることが分かる。これはつまり、再帰的推定手法 RENC の繰り返し回数 n が増加するにつれ、ノード特性の推定精度が向上していることを意味している。ネットワークのトポロジ構造をすべて利用するのではなく、ノード特性の推定にとって意味のあるノードやリンクのみを用いることにより、ノード特性の推定精度を向上できていると考えられる。また、再帰的推定を $n = 4 \sim 5$ 程度まで繰り返すことにより、ノード特性の推定精度が向上していることが分かる。

なお、紙面の都合上結果は省略するが、ノードの特性値 U_i をパレート分布によって与えた時も、同様な結果が得られていた。ただし、ノードの特性値の分布と再帰的推定手法 RENC の有効性との関係については、今後さらなる調査が必要である。

6 まとめと今後の課題

本稿では、ネットワークのトポロジ構造からノード特性を再帰的に推定する手法 RENC を提案した。また、ソーシャルネットワークのいくつかの特徴を再現することができる、ネットワーク生成モデル LRE を提案した。そして、LRE によって生成したネットワークに RENC を適用することにより、ノード特性の再帰的推定手法 RENC の有効性を評価した。その結果、再帰的推定手法 RENC により、ノード特性の推定精度が向上する

ことが分かった。

今後の課題として、再帰的推定手法 RENC の現実の複雑ネットワークを対象とした評価を行う予定である。また、RENC の最適な繰り返し回数を決定させる必要がある。さらに、再帰的推定手法を応用し、情報検索・情報分析サービス、コミュニティ支援サービスの考案などが挙げられる。

謝　　辞

本研究を実施するにあたり、有意義な議論をしていただいた、日本電信電話株式会社 NTT 情報流通プラットフォーム研究所の村山純一氏、田邉正雄氏、波戸邦夫氏、八木毅氏に感謝する。

文　　献

[1] R. Albert and A.-L. Barabasi, “Statistical mechanics of complex networks,” *Reviews of Modern Physics*, vol. 74, June 2002.

[2] M. Jin, M. Girvan, and M. E. J. Newman, “The structure of growing social networks,” *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 64, pp. 381–399, Oct. 2001.

[3] D. Liben-Nowell and J. Kleinverg, “The link prediction problem for social networks,” in *Proceedings of the 12th ACM International Conference on Information and Knowledge Management (CIKM’03)*, pp. 556–559, June 2003.

[4] G. F. Davis, M. Yoo, and W. E. Baker, “The small world of the american corporate elite, 1982-2001,” *Strategic Organization*, vol. 1, pp. 301–326, Aug. 2003.

[5] E. Holger, M. Lutz-Ingo, and B. Stefan, “Scale-free topology of e-mail networks,” *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 66, pp. 1–4, Sept. 2002.

[6] M. Granovetter, “The strength of weak ties: a network theory revisited,” *Sociological Theory*, vol. 1, pp. 201–233, Sept. 1983.

[7] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature (London)*, vol. 393, pp. 440–442, Oct. 1998.

[8] A. L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Viesek, “Evolution of the social network of scientific collaborations,” *Physica A*, vol. 311, pp. 590–614, Jan. 2002.

[9] H. A. Kautz, B. Selman, and M. A. Shah, “The hidden Web,” *AI Magazine*, vol. 18, no. 2, pp. 27–36, 1997.

[10] K. Sugiyama, O. Honda, H. Ohsaki, and M. Imase, “Application of network analysis techniques for Japanese corporate transaction network,” in *Proceedings of 6th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT 2005)*, no. 9–10, pp. 387–392, Nov. 2005.

[11] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the Web,” tech. rep., Stanford Digital Library Technologies Project, 1998.

[12] 安田 雪, ネットワーク分析 — 何が行為を決定するか. 新曜社, Feb. 1992.

[13] A. L. Barabsi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, pp. 509–512, Oct. 1999.

[14] A. Fronczak, P. Fronczak, and J. A. Holyst, “Average path length in random networks,” *Physical Review E*, vol. 70, Nov. 2004.

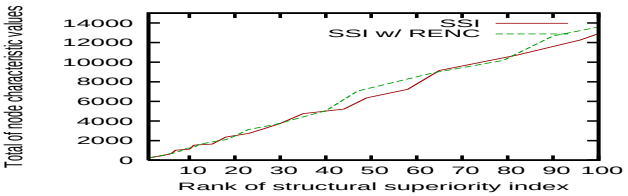

図 3: ノードの順位とノード特性値の累計の関係 ($n = 1$ の場合)

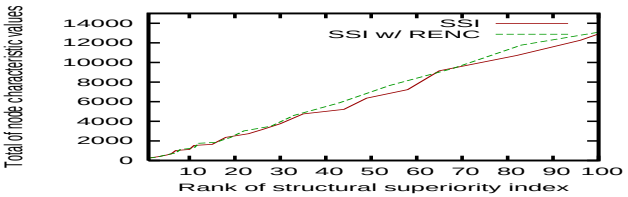

図 4: ノードの順位とノード特性値の累計の関係 ($n = 2$ の場合)

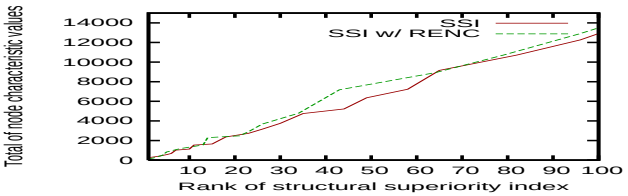

図 5: ノードの順位とノード特性値の累計の関係 ($n = 3$ の場合)

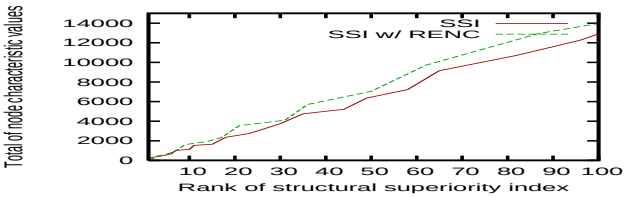

図 6: ノードの順位とノード特性値の累計の関係 ($n = 4$ の場合)

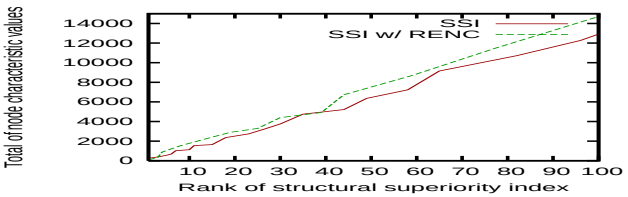

図 7: ノードの順位とノード特性値の累計の関係 ($n = 5$ の場合)

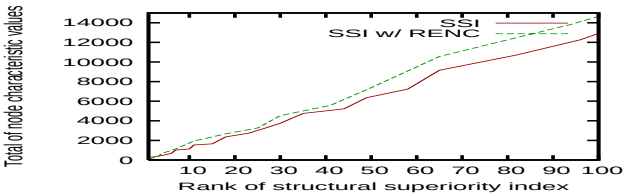

図 8: ノードの順位とノード特性値の累計の関係 ($n = 6$ の場合)

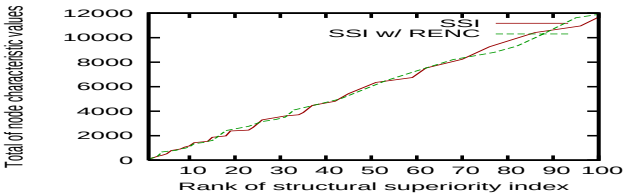

図 9: ノードの順位とノード特性値の累計の関係 ($n = 1$ の場合)

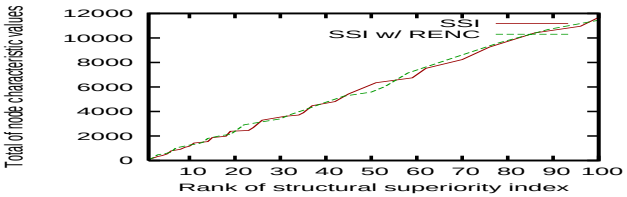

図 10: ノードの順位とノード特性値の累計の関係 ($n = 2$ の場合)

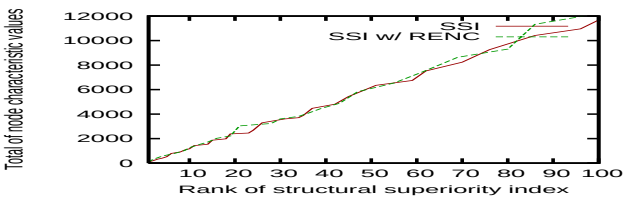

図 11: ノードの順位とノード特性値の累計の関係 ($n = 3$ の場合)

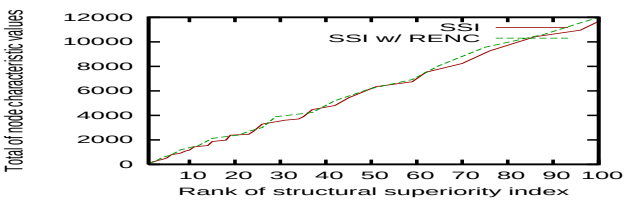

図 12: ノードの順位とノード特性値の累計の関係 ($n = 4$ の場合)

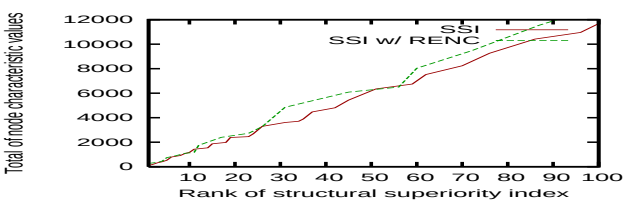

図 13: ノードの順位とノード特性値の累計の関係 ($n = 5$ の場合)

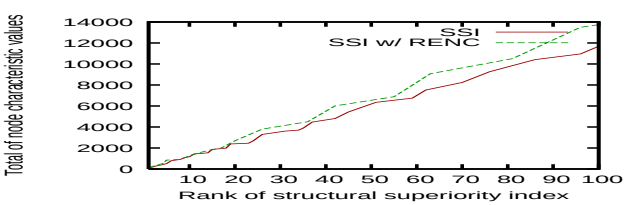

図 14: ノードの順位とノード特性値の累計の関係 ($n = 6$ の場合)