

# セマンティックグラフモデルを用いたステークホルダ分析方法の提案と評価

白崎 悠太<sup>†1</sup> 小林 勇也<sup>†1</sup> 青山 幹雄<sup>†1</sup>

**概要:** システム開発において、ステークホルダ分析の重要性が高まっている。本研究では、セマンティックグラフモデルを用いたステークホルダ分析方法を提案する。システム開発会議の議事録データに対して形態素解析と係り受け解析を行うことで、ステークホルダと目的語との関係を抽出し、セマンティックグラフモデルを生成する。グラフデータベースを用いて、生成したセマンティックグラフモデルを分析し、ステークホルダの構造特性を明らかにし、可視化する方法を提案する。提案方法を実際の開発議事録、約 27,000 字に適用し、提案方法の有効性を示す。

**キーワード:** 要求工学, ステークホルダ分析, セマンティックグラフモデル, グラフデータベース, グラフ分析, データ駆動

## A Stakeholder Analysis Method Based on Semantic Graph Models

YUTA SHIRASAKI<sup>†1</sup> YUYA KOBAYASHI<sup>†1</sup> MIKIO AOYAMA<sup>†1</sup>

### 1. はじめに

近年、ソフトウェア要求獲得における要求仕様の高品質化が重要になっている。そのため、要求獲得の源流であるステークホルダ分析の重要性が高まっている[4,10]。しかし、ステークホルダ分析で必要となるインタビューやアンケートなどを通して得られる情報は多様で膨大である。よって、人手に頼っている現状では、十分にデータを活用した合理的なステークホルダ分析は困難である。

本研究では、データに基づく高品質な要求獲得を支援するステークホルダ分析方法の確立を課題とする。要求獲得で得られるデータのステークホルダ間の関係に着目し、セマンティックグラフモデルで表現することで、ステークホルダの意味分析を行う方法を提案する。また、セマンティックグラフモデルのデータベースとして、グラフデータベースを採用する。

本研究では、以下の 3 点を研究課題とする。

- (1) テキストデータを日本語テキスト解析技術によって分析し、セマンティックグラフモデルでモデル化する方法の提案
- (2) グラフデータベース上での発話者ステークホルダの意味分析方法の提案
- (3) 実データへの適用による、提案方法の妥当性の確認

### 2. 関連研究

#### 2.1 ステークホルダ分析

ステークホルダ分析は要求定義において最初に行う作業であり、要求の品質の安定に繋がる工程である。要求を獲得する源泉となるステークホルダの識別方法を開発時に取り入れ、分析した結果を可視化し、それらを元に関係者で合意し、合意されたステークホルダを中心に要求定義することが重要である。

#### 2.2 データ駆動要求獲得

関連研究において、データに基づく要求獲得方法の提案と発話者ステークホルダとグラフの測地線距離に基づく話題ステークホルダの構造化方法の提案がされている[2]。また、この提案によって得られる期待効果として、事実に基づく合理的な要求獲得が可能になること、意味的構造を含むデータ解析が可能になることの 2 点を挙げている。一方で、課題としてセマンティックグラフモデルを利用した意味分析方法が未確立であるといえる。

#### 2.3 セマンティックグラフモデル

セマンティックグラフモデルとは、グラフ構造を持ち、格納しているデータそのものだけでなく、従来では表現が困難であったデータの相互関係を表現可能なデータモデルである[7,9]。セマンティックグラフモデルにおける最小単位はノードであり、人やデータ等の調査する対象を表す。また、ノード間のエッジをリレーションシップという。ノード間の関連を表し、始点から終点への向きを持つ。さらにリレーションシップに対して、ノード間の関係をタイプとして定義できる。本研究ではセマンティックグラフモデルの中でもノード、リレーションシップの属性をプロパティとして付与できるプロパティグラフモデルを採用する。

#### 2.4 Neo4j

本研究では、セマンティックグラフモデルのデータベースとして、NoSQL のグラフ型データベースとして広く用いられている Neo4j を用いる[8]。Neo4j の特性を以下に示す。

- (1) スケーラビリティである。
- (2) 宣言型クエリ言語 Cypher を用いている。
- (3) 幅広い開発言語やフレームワークに対応している。
- (4) オープンソースとして提供されている。

#### 2.5 要求獲得における日本語テキスト解析技術

<sup>†1</sup> 南山大学 理工学部 ソフトウェア工学科  
Dep. of Software Engineering, Nanzan University

テキスト解析とは、通常の文章からなるデータを単語や文節で区切り、それらの出現頻度や出現傾向、時系列などを解析することで有用な情報を抽出する、テキストデータの分析方法である[1, 5]。日本語文章の構造化解析として、論説文などの文章を書き手の主張という観点から構造化する方法が提案されている[3]。

さらに、会議の発話履歴を構造化し、それを基に構造化された仕様書を作成する方法が提案されている。会議の記録から直接的に仕様書を作成するため、会議で議論されながらも仕様書には記述が残らないような欠落事項の発生の防止が期待効果として挙げられている。また、会議毎に記録を構造化することで、効率的な議論の促進、不明瞭な事項の発生の防止などが期待できる。一方で課題として、作業者の主観に依存して構造化をするという問題点がある[6]。

### 3. アプローチ

#### 3.1 ステークホルダ分析の定義

本研究におけるステークホルダ分析の対象範囲を以下のように定義する。

- (1) 対象とするステークホルダ
  - a) 会議に参加し、1回以上何らかの発言をした発話者
  - b) 発話の中で登場した人物
- (2) 目的と分析対象
  - a) 議事録の発話データからステークホルダの位置付けを分析することを目的とする。
  - b) あらかじめ分かっている役職などのステークホルダの情報は、分析対象とはしない。

#### 3.2 ステークホルダ分析のアプローチ

アプローチを図1に示す。



図1 アプローチ  
 Figure 1 Approach

本研究では、議事録データからテキスト解析をして得られたデータに対し、ステークホルダ間の関係の意味付けが可能になるセマンティックグラフモデルを用いて、ステークホルダ間の関係を分析し、そのステークホルダの構造を可視化することによって、要求獲得を可能とする。ステークホルダ分析として、グラフ分析と統計的な構造分析を行う。可視化した分析結果からデータに基づく要求獲得を支援可能であると考えられる。

### 4. 提案方法

#### 4.1 提案プロセス

本研究では、セマンティックグラフモデルを用いたステークホルダ分析方法を提案する。各プロセスとそれらによって得られる成果物の流れを表した提案プロセスを図2に示す。

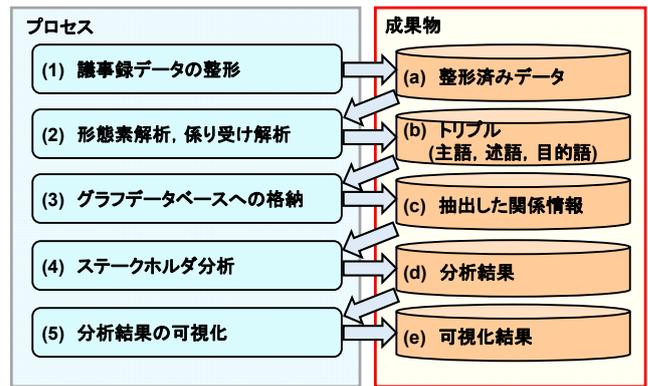


図2 提案プロセス  
 Figure 2 Proposed process

提案プロセスの詳細は以下の通りである。

- (1) 議事録データの整形  
 議事録データを、形態素解析と係り受け解析を行える形式にテキストデータとして整形する。
- (2) 形態素解析, 係り受け解析  
 整形済みデータの各文に対して MeCab[5], CaboCha[1]を適用し、形態素解析, 係り受け解析を行い、トリプル(主語, 述語, 目的語)を抽出する。
- (3) グラフデータベースへの格納
  - a) セマンティックグラフモデルの生成  
 (2)で抽出したトリプルからセマンティックグラフモデルを生成する。扱う議事録データによって、ノードとリレーションがプロパティとして持つデータを検討する。
  - b) 格納処理  
 生成したセマンティックグラフモデルに従い、Neo4jのクエリ言語である Cypher を用いてグラフデータベースにデータを格納する。
- (4) ステークホルダ分析
  - a) グラフ分析  
 Cypherを用いて1つの議題についてのグラフを抽出し、分析を行う。分析は大域的グラフ分析と局所的グラフ分析の2段階分析を行う。
  - b) ステークホルダマトリクスによる分析  
 グラフ分析の結果を受けて、ステークホルダの関与度と影響度からステークホルダマトリクスを生成し、分析する。
- (5) 分析結果の可視化  
 分析結果に基づき、Neo4jのデータを更新し、ステークホルダ間の関係を可視化する。

#### 4.2 分析方法

ステークホルダの構造分析方法として2つの方法を提案する。

##### 4.2.1 グラフ分析

- (1) 大域的グラフ分析  
 大域的グラフ分析を、セマンティックグラフモデルとして表現した議事録データ全体と特定の議題全体のグラフに対して行う分析として定義する。グラフの形状、各ノードの度数に基づく議題の中心人物を分析する。

ノードの次数とは、あるノードと他のノードとの関係数である。有向グラフの場合、ノードの次数は、出次数と入次数と 2 つの尺度が存在する。出次数を基準とした中心性は、他者への働きかけとしての中心性の程度を考える上で有効である。入次数を基準とした中心性は、グラフ内の他者の行為の対象となる中心性である。ここでは、発話者ステークホルダを分析対象とするので発話者の発言数をノードの次数として分析する。

(2) 局所的グラフ分析

局所的グラフ分析を、特定のステークホルダ集合の意味構造分析として定義する。ステークホルダ間の関係に着目するために大域的グラフ分析で明らかになった複数のステークホルダと関係を持つ目的語を含むサブグラフを出力し、分析をする。

サブグラフとは、大域的分析で用いたグラフに含まれる一部のノードとリレーションシップの集合を指す。グラフを構成するノードとリレーションシップのうち、一部のノードとリレーションシップによって形成される部分グラフである。

局所的グラフ分析の適用範囲例を図 3 に示す。実線で囲まれた範囲のグラフを分析対象のクリークとして扱う。のグラフ全体を大域的グラフ分析で扱う。

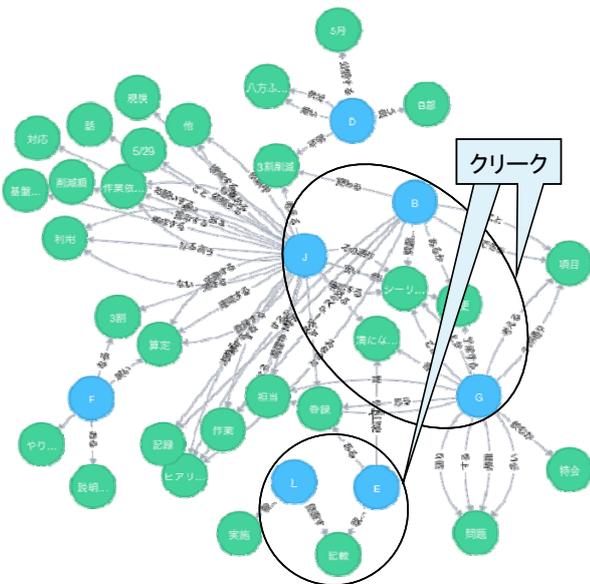


図 3 大域的、局所的の範囲例

Figure 3 An Example of Global and Local Graph

クリークとは、グラフ内で直接連結し、相互に強い関係で結ばれている複数のステークホルダの集合である[11]。ここでは、目的語ノードからの測地線距離が 1 のノードの集合を分析対象のクリークとして定義する。

局所的グラフ分析では、特定したクリークを分析する。

a) 単項ステークホルダ分析

1 人の特定のステークホルダの関心とその強さを目的語への関係の意味と関係数によって分析し、そのステークホルダの役割を推定する。

b) 二項ステークホルダ分析

単項分析で明らかになったステークホルダの役割を前提とし、グラフから、ある目的語を介して直接リレーションシップで接続

される関係を持つ 2 人のステークホルダの集合(クリーク)をサブグラフとして抽出し、分析する。図 4 に二項関係の例を示す。二項関係から、ステークホルダ間の協調や競合といった相互関係と従属関係を分析する。

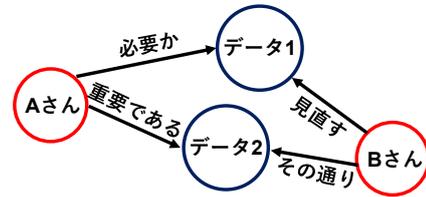


図 4 二項関係の例

Figure 4 An Example of Binary Relationship

c) 三項ステークホルダ分析

二項分析と同様に、単項分析で明らかになったステークホルダの役割を前提とし、グラフから、ある目的語を介して直接リレーションシップで接続されるトライアド(三者関係)を持つ 3 人のステークホルダの集合(クリーク)をサブグラフとして抽出し、分析を行う。図 5 に三項関係の例を示す。三項関係から、1 対多のステークホルダ間の相互関係を分析する。同じ役割を担っているステークホルダの集合を特定する。

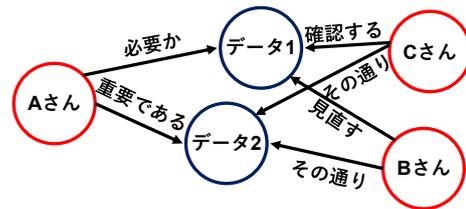


図 5 三項関係の例

Figure 5 An Example of Ternary Relationship

4.2.2 ステークホルダマトリクスによる分析

局所的グラフ分析で明らかになったステークホルダの意味構造を、統計的に分析する。議題の中心人物と関係を持つステークホルダの構造を明らかにする。

関与度、影響度をそれぞれ式(1)、式(2)で定義する。

$$\text{関与度} = \frac{\text{各ステークホルダの発言数}}{\text{議題に関するステークホルダ全体の合計発言数}} \quad (1)$$

$$\text{影響度} = \frac{\sum(\text{発話意図の重み} \times \text{各発話意図の発言数})}{\text{各ステークホルダの発言数}} \quad (2)$$

関与度は、そのステークホルダがどの程度議題に関わっているかの指標である。影響度は、そのステークホルダが他のステークホルダに対して、どの程度の影響力を持っているかの指標である。

関与度と影響度をそれぞれ縦軸と横軸にとり、ステークホルダマトリクスを生成する。生成したステークホルダマトリクスからクラスを特定する。グラフ分析の結果を受けて、特定したクラスの意味付けを行う。

## 5. ステークホルダ分析システムのプロトタイプ

議事録データは膨大であり、手作業でのテキスト解析とグラフデータベースへの格納処理は困難である。よって、図 6 に示すプロトタイプを実装し、一連の処理を自動化することにより、テキスト解析、グラフ分析、構造分析という 3 つのプロセスを実現した。

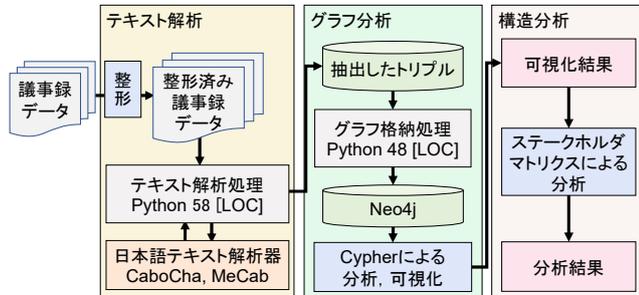


図 6 システム構成図

Figure 6 System Configuration Diagram

- (1) テキスト解析プロセス  
 テキスト解析プロセスにおけるテキスト解析処理では、Python を用い、58 行で記述した。
- (2) グラフ分析プロセス  
 グラフ分析プロセスにおけるグラフ格納処理でも同様に Python を用い、48 行で記述した。

## 6. 議事録データへの適用

### 6.1 目的と適用対象

本研究での提案プロセスを実際の議事録に適用、評価し、提案プロセスの妥当性を示す。適用対象は、公共情報システム開発の約 27,000 字で構成される実際の議事録データである。以下の表 1 に各会議における文章量と実施日を示す。

表 1 各会議における文章量と日付

Table 1 Amount of Document and Date at Each Meeting

会議	文章量(文字)	日付
第1回	6125(文字)	2015/4/21
第2回	6015(文字)	2015/4/28
第3回	4620(文字)	2015/5/12
第4回	7577(文字)	2015/5/19
第5回	7311(文字)	2015/5/26

### 6.2 議事録データの整形

本研究で取り扱う議事録データには、表形式のデータ、タイトル、会議の実施日など、そのままの形式では形態素解析と係り受け解析に適さないデータと不必要なデータが含まれている。したがって、議事録データを形態素解析、係り受け解析に適した形式に整形した。

### 6.3 形態素解析、係り受け解析

MeCab と CaboCha を用いて、整形済みデータを形態素解析、係り受け解析することで、主語、述語、目的語のトリプルを抽出する。下記にプロセスを示す。

- (1) 係り先のない文節から述語を抽出する。

(2) 適用対象の議事録データには、(発話者)のように、発話者が括弧で囲まれた形で書かれているので、括弧内を主語として抽出する。

- (3) 主節の係り先の文節の名詞を目的語として抽出する。

## 6.4 グラフデータベースへの格納

### 6.4.1 データモデル

Neo4j のクエリ言語である Cypher を用いて、図 7 のデータモデルに基づき、データを Neo4j に格納する。

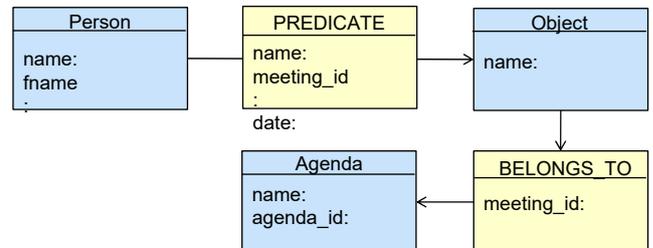


図 7 データモデル

Figure 7 Data Model

- (1) ノード  
 図中の青色のクラスは、ノードを表している。以下に各ノードの概要を示す。
  - a) Person  
 会議における、発話者を表している。プロパティとして、発話者の氏名 (name) と仮名 (fname) を持つ。
  - b) Object  
 発話者が発した発話文中の目的語を表している。プロパティとして、目的語の名前 (name) を持つ。
  - c) Agenda  
 会議での議題を表している。プロパティとして、議題の議事名 (name) と議事録番号 (agenda\_id) を持つ。
- (2) リレーションシップ  
 黄色のクラスは、リレーションシップを表している。以下に各リレーションシップの概要を示す。
  - a) PREDICATE  
 発話者と目的語を結ぶ、発話者が発した発話文中の述語を表している。プロパティとして、述語名 (name)、会議番号 (meeting\_id)、会議の日付 (date) を持つ。
  - b) BELONGS\_TO  
 目的語と議題を結ぶ、目的語がどの議題に属しているかのリレーションシップを表している。プロパティとして、会議番号 (meeting\_id) を持つ。

### 6.4.2 格納プロセス

生成したデータモデルに基づき、ノードとリレーションシップをグラフデータベースへ格納し、可視化した結果を図 8 に示す。

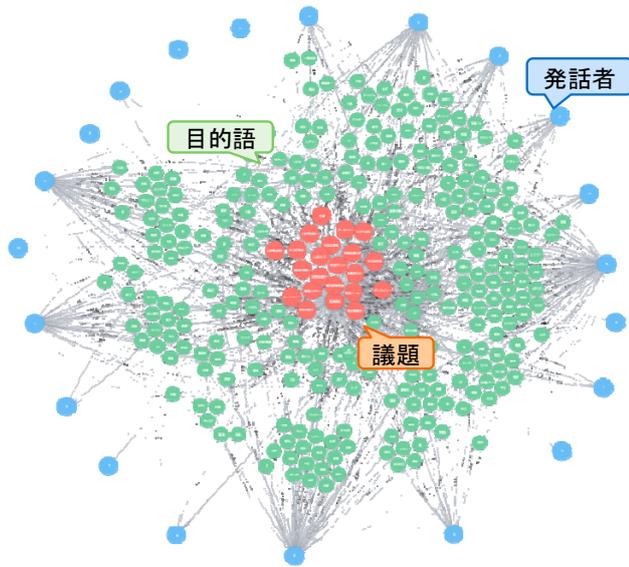


図 8 リレーションシップ追加後のグラフ  
 Figure 8 AGraph after Adding Relationship

グラフデータベースへの格納プロセスを以下に示す。

- (1) 主語ノードの格納  
 発話者を主語ノードとして格納する。格納したそれぞれの主語ノードをステークホルダとして、本研究での分析対象とする。ここで、発話者には同じ名前の者は存在しないと仮定する。
- (2) 目的語ノードの格納  
 目的語ノードも主語ノードと同様に、重複を許さず格納する。
- (3) 議題ノードの格納  
 議題ノードもまた同様に、重複を許さず格納する。
- (4) リレーションシップの追加  
 存在するノード間のリレーションシップを追加する。

### 6.5 ステークホルダ分析

格納したデータに対して、Cypher を用いてステークホルダ分析を行う。図に示した通り、議事録データには多数の議題ノードが存在し、議題ごとに関わるステークホルダは異なる。本研究では、会議において多くの発言があり、議論が活発に行われ、重要性が高いと推定される特定の議題に焦点を絞って、ステークホルダの構造を明らかにする。

#### 6.5.1 グラフ出力

「議題 No.8 データベース登録支援」では、多くの発言があり、議論が活発に行われたことがわかる。よって、「議題 No.8 データベース登録支援」に関わる主語ノードと目的語ノードを含むグラフを出力する。

#### 6.5.2 グラフ分析

本研究では、グラフ分析の方法として、グラフ全体の構造分析をする大域的グラフ分析とそれぞれのステークホルダと目的語との関係からステークホルダの構造を分析する局所的グラフ分析の2段階分析を提案する。

### (1) 大域的グラフ分析

図 9 に「議題 No.8 データベース登録支援」のグラフを示す。図 9 より、主語ノードを円状に均等に配置しているにもかかわらず、目的語ノードが右に偏っている。その影響で、議題ノードもグラフの中心より右寄りであり、グラフの重心がずれていると言える。これは、他のステークホルダよりも発言数が多い主語ノードが目的語ノードを引きつけているため、引き起こされる。そこから、グラフの右側のステークホルダと目的語との間での議論が活発に行われていることが分かる。

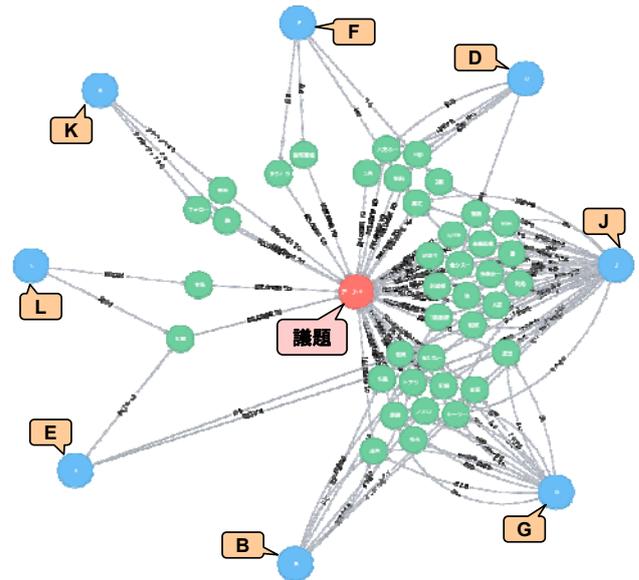


図 9 議題 No.8 データベース登録支援  
 Figure 9 Database Registration Support for Agenda No.8

「議題 No.8 データベース登録支援」についての発言数分布を図 10 に示す。グラフの右側に位置している J, G, B の発言数が多いことが分かる。特に J の発言数が多く、この議題に対する中心人物として推定される。

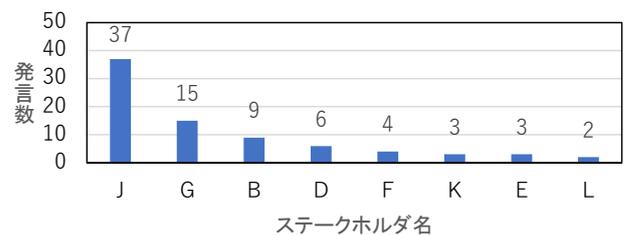


図 10 ステークホルダの発言数分布  
 Figure 10 Number of Remarks per Stakeholder

### (2) 局所的グラフ分析

ステークホルダ間の関係に着目するため、大域的グラフ分析で明らかになった、複数のステークホルダと関係を持つ目的語を含むサブグラフを出力する。図 11 に示す。

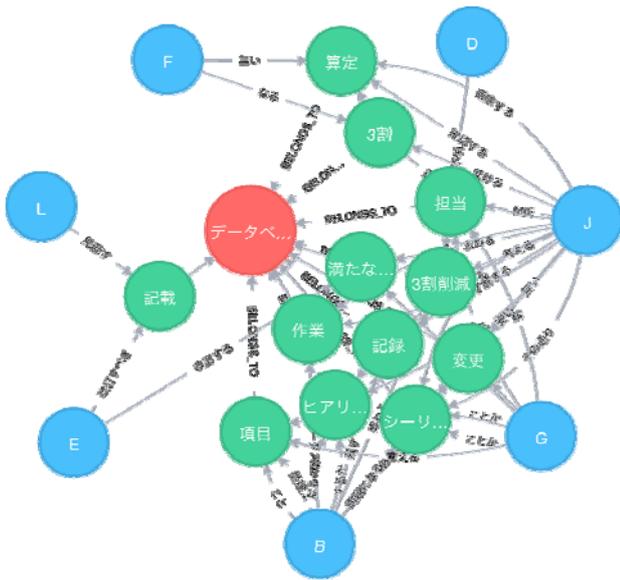


図 11 抽出したサブグラフ  
 Figure 11 Extracted Subgraph

抽出したサブグラフ上の単項、二項関係、三項関係に着目し、ステークホルダの役割を分析する。

- a) 単項ステークホルダ分析
- i) Kの役割

Kは、図10に含まれているが、図11には含まれていない。したがって、他のステークホルダと共通の目的語を持たず、他のステークホルダ全体に対して一方的な報告を行なっていることが分かる。

- ii) Jの役割

図9と図11を比較すると、Jの関係数が著しく減少している。また、Jの目的語への関係を見ると、報告と質問への返答が多い。よって、Jは議題に対して深く関係しており、他のステークホルダに情報を与える立場であることが分かる。

- b) 二項ステークホルダ分析

単項分析で明らかになったステークホルダの役割を前提として分析を行う。

- i) EとLの関係

EとLの関係に注目する。EとLの二項関係を図12に示す。



図 12 EとLの二項関係  
 Figure 12 Binary Relationship of E and L

Eの「改訂すべき」という要望に対して、Lが「見直す」と返答しており、EがLを監督する立場であることが分かる。

- ii) DとJ, FとJの関係

DとJ, FとJの関係に注目する。DとJ, FとJの二項関係を図13に示す。

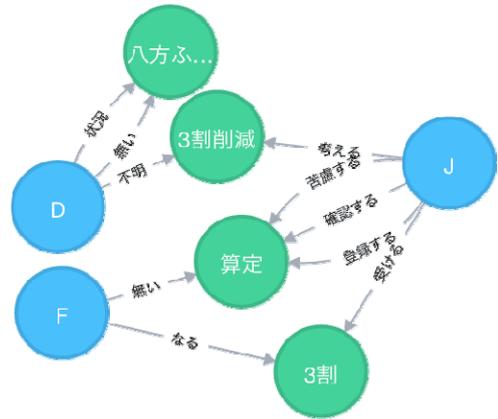


図 13 DとJ, FとJの二項関係  
 Figure 13 Binary Relationship of D and J, F and J

Dに関して、Jが目的語である「3割削減」に対して「考える」と報告をし、Dが「不明」と意見を言っていることが分かる。

また、Fに関して、Jが目的語である「算定」に対して「確認する」、「登録する」等の報告をし、Fが「無い」と意見を言っていることが分かる。

よって、両者の発話ともにJが情報を与えた目的語への意見であることから、議題自体には深く関係していないが、特定の話題について意見する立場であることが分かる。

- c) 三項ステークホルダ分析

B, G, Jに関係を持つ共通の目的語が存在する。したがって、その部分をサブグラフとして抽出し、三項関係を分析する。B, G, Jの三項関係を図14に示す。

単項分析で明らかになったJの役割を前提として分析を行う。「変更」に対してBが「あるか」と質問すると、Jが「無い」と返答し、Gが「了承する」と返答している。また、「シーリング額」に対してBが「問題になる」と問題提起をし、GがJに「ことか」と質問し、Jが「その通り」、「確認する」と返答している。

よって、BとGはJよりも上位の監督する立場のステークホルダである。また、そこからBとGは、議題に対する影響力が強いステークホルダであることが分かる。

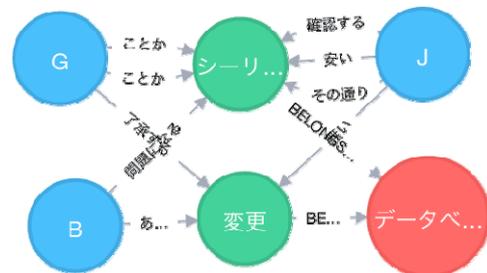


図 14 B, G, Jの三項関係  
 Figure 14 Triadic Interaction of B, G and J

### 6.6 分析結果の可視化

2段階分析の結果を踏まえ、報告を行う関係には REPORTS\_TO, 管理する関係には MANAGES, 意見や助言

を行う関係には OPINION\_TO の関係を追加した。可視化した発話者ステークホルダ間の関係を図 15 に示す。

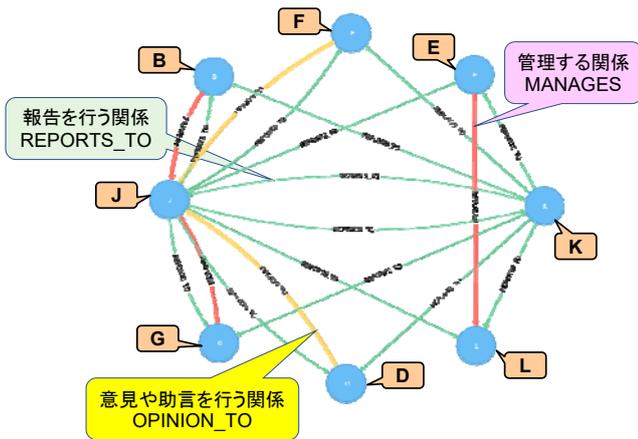


図 15 ステークホルダ間の関係

Figure 15 Relationships between Stakeholders

### 6.7 ステークホルダマトリクスによる分析

#### 6.7.1 分析対象

グラフ分析の結果を受けて、主たる発言者である J を中心としたステークホルダの関係を定量的に分析する。E, L, K の 3 者と J との関係は、会議における一方的な報告の関係のみであるので、ここでは分析の対象とはしない。

#### 6.7.2 発話意図の定義と重み付け

MANAGES に属する発言は、問題提起など他のステークホルダの行動を管理する発言とし、OPINION\_TO に属する発言は、話題についての評価、意見を述べる発言とする。発話意図の重みを定義し、表 2 に示す。

表 2 発話意図の重み

Table 2 Weight of Speech Intention

	MANAGES	OPINION_TO
	0.8	0.2

#### 6.7.3 関与度と影響度の集計

ステークホルダに対する発話意図の発言数を集計し、関与度と影響度を計算した結果を表 3 に示す。

表 3 各ステークホルダの発言集計結果

Table 3 Sum of Statements of Each Stakeholder

	MANAGES	OPINION_TO	関与度	影響度
B	6	1	0.13	0.44
D	0	3	0.08	0.2
F	0	2	0.06	0.2
G	4	5	0.21	0.29
J	0	2	0.52	0.01

表 3 において、MANAGES, OPINION\_TO の列で各ステークホルダのそれぞれの発言数、関与度、影響度の列で各ステークホルダの関与度、影響度の値を示す。

#### 6.7.4 分析結果の可視化とクラスタの特定

各ステークホルダの発言集計結果より、得られたステークホルダの影響度と関与度を図 16 に示すステークホルダマトリクス上に表現することで、下記のように 3 つのクラスタを特定した。さらに、グラフの意味分析の結果を受けて、クラスタに含まれているステークホルダの役割と関与度、影響度の値からクラスタの意味付けを行った。

##### a) 管理者クラスタ

管理者クラスタには、他のステークホルダを管理するステークホルダが含まれる。ここで、管理者クラスタに分類されるステークホルダは、影響度が大きく、関与度が小さい。

よって、管理者クラスタに分類されるステークホルダは、他のステークホルダをまとめる上位の立場の人ステークホルダであると推定される。

##### b) 御意見番クラスタ

御意見番クラスタには、特定の話題について意見、助言を行うステークホルダが含まれる。ここで、御意見番クラスタに分類されるステークホルダは、影響度が小さく、関与度が小さい。

よって、御意見番クラスタに分類されるステークホルダの会議における役割は、社外、もしくは他の部署に所属し、開発するシステムに関する技術に精通する知識人であると推定される。

##### c) 報告者クラスタ

報告者クラスタには、他のステークホルダに対して、議題についての報告を行うステークホルダが含まれる。ここで、報告者クラスタに分類されるステークホルダは、影響度が小さく、関与度が大きい。

よって、報告者クラスタに分類されるステークホルダの会議における役割は、上位の立場のステークホルダに対して報告を行うので、会議における下位の立場のステークホルダであると推定される。

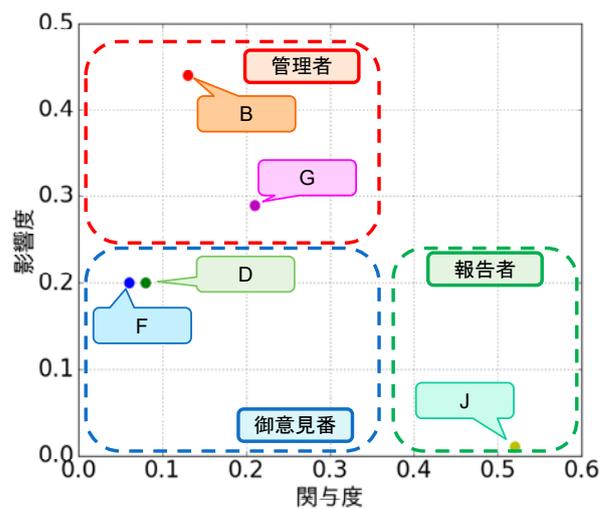


図 16 ステークホルダマトリクス

Figure 16 Stakeholder Matrix

## 7. 評価

### 7.1 セマンティックグラフモデルによるモデル化評価

セマンティックグラフモデルを用いることで、従来のデータモデルでは表現が困難であった要素間の意味的關係が表現可能となった。さらに、要素間の意味的關係を分析することで、ステークホルダ間の意図を明らかにすることが可能となった。

### 7.2 グラフデータベース上での発話者ステークホルダの意味分析方法の提案の評価

発話者ステークホルダに対して大域的グラフ分析と局所的グラフ分析の2段階分析方法を提案した。大域的分析では、グラフの中心性から、議題についての中心人物を分析することが可能となった。共通の目的語を持つステークホルダ間の関係に着目した意味分析により、各ステークホルダの役割を推定可能であることを示した。それによって、ステークホルダの構造化が可能になった。さらに、グラフ分析によってステークホルダの役割を特定し、発話データに基づき影響度、関与度を定量的に評価した。この結果をステークホルダマトリクス上にマッピングすることで、ステークホルダをクラスタ化し、その役割を推定できることを示した。

### 7.3 実データへの適用による提案方法の妥当性の評価

本研究では、会議5回分の議事録データに対して本提案方法を適用し、提案方法の妥当性を検証した。会議を重ねる度に議事録データ量が増加すると考えられるが、グラフデータベースを用いることでデータモデルを定義し直すことなく、追加の議事録データの格納と分析が行える拡張性があり、かつ、同一モデル上で議事の変化を分析することができる。

しかし、テキストデータに対して日本語テキスト解析を行い、グラフデータベースに格納する方法の提案については、形態素解析の精度の問題があり、手作業での修正が必要であるという問題が生じた。

## 8. 考察

### 8.1 従来のステークホルダ分析方法との比較

従来のステークホルダマトリクスを用いた分析方法[10]では、分析が困難なステークホルダ間の構造をデータに基づき分析可能という点で、より有効性が高い。

また、分析後に構造化したステークホルダ間の関係をセマンティックグラフモデルで可視化することで、要求獲得者がステークホルダの関係を視覚的に把握することができる。したがって、可視化結果から、要求アナリストの能力に依存しない要求獲得が可能となり、要求獲得の効率化と要求仕様の高品質化が期待できる。

### 8.2 セマンティックグラフモデルを用いたステークホルダ分析方法との比較

セマンティックグラフモデルを用いたステークホルダ分析方法が提案[2]されているが、本研究では、新たに大域的グラフ分析と局所的グラフ分析の提案をした。グラフの意味分析を行うことで、ステークホルダマトリクスで特定したクラスタに対して、詳細な意味付けが可能になった。

### 8.3 セマンティックグラフモデルの適用範囲

Sharp らの提案したステークホルダ識別方法[10]での、ベースラインステークホルダ、サテライトステークホルダ等の分類は人手に頼っている。しかし、本提案方法を適用すれば、データに基づくグラフ分析が適用可能である。

## 9. 今後の課題

今後、下記の課題を検討する必要がある。

#### (1) セマンティックグラフモデルの応用

セマンティックグラフモデルの意味構造を利用した、より詳細な、高度な分析を検討する。また、データの適用範囲を拡大し、提案方法の妥当性を確認する必要がある。

#### (2) 日本語テキスト解析の改善

同一の意味を表す代名詞がノードとして存在してしまい、手直しの手間が生じた。よって、代名詞が指す名詞を一元管理し、同一名詞の割り当て処理を可能にする。

## 10. まとめ

本研究では、テキストデータを日本語テキスト解析技術によって分析し、セマンティックグラフモデルでモデル化する方法の提案とグラフデータベースに格納したテキストデータから、ステークホルダ、目的語、目的語に対する述語に着目し、発話者ステークホルダの意味分析を行う方法の提案をした。

実際のプロジェクト議事録に本提案方法を適用し、提案方法の有効性、妥当性を示した。

本提案を行うことで、従来のデータモデルでは表現することができない、ステークホルダの関係に着目した構造分析が可能となった。したがって、データに基づく高品質な要求獲得につながる支援が可能となる。

## 参考文献

- [1] CaboCha, Yet Another Japanese Dependency Structure Analyzer, <http://taku910.github.io/cabocha/>.
- [2] 藤本 玲子, 他, セマンティックグラフモデルによるデータ駆動要求工学の提案とステークホルダ分析への適用評価, SES 2016 論文集, Sep. 2016, pp. 179-186.
- [3] 福本 淳一, 他, 日本語文章の構造化解析, 自然言語処理 85-11, Sep. 1991, pp. 81-88.
- [4] 位野木 万里, 要求獲得におけるステークホルダ識別手法の実適用評価, 情報処理学会デジタルプラクティス, Vol. 4, No. 4, Apr. 2013, pp. 152-160.
- [5] MeCab, Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/>.
- [6] 三浦 信幸, 他, 仕様作成会議の発話履歴を用いて仕様書を作成する方法, 社団法人 電子情報通信学会, Jan. 1994, pp. 9-16.
- [7] M. Lal, Neo4j Graph Data Modeling, Packt Pub., 2015.
- [8] Neo Technology, neo4j, 2016, <http://neo4j.com/>.
- [9] I. Robinson, et. al., Graph Databases, O'Reilly, 2015.
- [10] H. Sharp, et. al., Stakeholder Identification in the Requirements Engineering Process, Proc of DEXA, 1999, pp. 387-391.
- [11] 安田 雪, 実践ネットワーク分析: 関係を解く理論と技法, 新曜社, 2001.