

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^{†1} 小林 勇也^{†1} 青山 幹雄^{†2}

概要: 要求獲得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、ステークホルダ相互とその対象との関係を表現できるセマンティックグラフモデルとして、ステークホルダインテンショングラフ(SIG)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、SIGを生成する。さらに、グラフデータベースを用いて、生成したSIGを分析し、ステークホルダの構造特性を明らかにして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^{†1} YUYA KOBAYASHI^{†1} MIKIO AOYAMA^{†2}

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品質向上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[3,10]。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。さらに、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。

2. 研究課題

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

3. 関連研究

3.1 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、要求の品質向上に繋がる工程である[4,10]。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[3]。

3.2 データ駆動要求獲得

要求獲得においてデータを活用する方法が提案され、さらに、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、この提案の期待効果として、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。しかし、一方で、セマンティックグラフモデルを利用したステークホルダの意味分析方法が未確立という課題がある。

3.3 セマンティックグラフモデル(SGM)

セマンティックグラフモデルとは、グラフ構造上で、従来は表現が困難であったデータ間の意味定義を表現可能とするモデルである[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、人やもの等の対象を表す。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[9]を採用する。

セマンティックグラフモデルを扱うデータベースとして、グラフデータベースがある。実装として Neo4j[8]や Titan などがある。

3.4 要求獲得における日本語テキスト解析技術

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[7]。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。しかし、構造化を作業者の主観に依存する課題もある。

^{†1} 南山大学大学院 理工学研究科 ソフトウェア工学専攻
Graduate Program of Software Engineering, Nanzan University

^{†2} 南山大学 理工学部 ソフトウェア工学科
Dep. of Software Engineering, Nanzan University

4. アプローチ

4.1 ステークホルダ分析の範囲

本研究でのステークホルダ分析の対象範囲を以下に示す。

- (1) 対象とするステークホルダ
 - a) 会議に参加し、1回以上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、分析に利用しない。

4.2 ステークホルダ分析のアプローチ

ステークホルダ分析のアプローチを図1に示す。

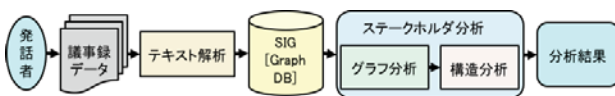


図1 アプローチ

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より、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生成する。ステークホルダ間の関係を分析し、構造を可視化する。ステークホルダ分析として、グラフ分析と定量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える。

5. ステークホルダ分析方法

5.1 ステークホルダ分析プロセス

提案する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各プロセスとそれらによって得られる成果物の流れを図2に示す。

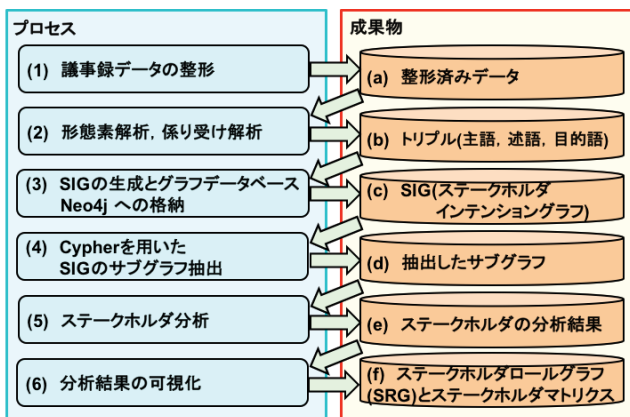


図2 提案プロセス

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

議事録データから会議の題名、日付などの処理に適さない情報を取り除き、形態素解析と係り受け解析を行える形式に議事録データを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[6], 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) SIGの生成とグラフデータベース Neo4j への格納

- a) SIGの生成

(2)で抽出したトリプルからSIG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) グラフデータベース Neo4j への格納

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 Cypher を用いてグラフデータベースにデータを格納する。

- (4) Cypherを用いたSIGのサブグラフ抽出

議事録データには、複数の議題が含まれている。議題毎に、関係するステークホルダとその役割は異なり、分析は議題毎に行う必要がある。したがって、ステークホルダ分析では、1つの議題に関係するステークホルダと目的語を含むSIGのサブグラフを分析対象とする。Cypherを用いて分析対象のサブグラフを抽出する。

- (5) ステークホルダ分析

本研究でのステークホルダ分析では、グラフ分析とステークホルダマトリクスを用いた分析を行う。

- 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の生成と分析

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。

- (6) 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。分析結果の可視化結果として、SRGとステークホルダマトリクスが得られる。

5.2 グラフモデル

5.2.1 ステークホルダ分析のグラフモデル

図3にデータ間の関係を示す。整形済み議事録データからSIGが生成される。SIGは、分析対象のグラフであり、Cypherによるステークホルダ分析を行い、分析結果としてSRGが生成される。SIGの定量的分析からステークホルダマトリクスが生成される。生成されたステークホルダマトリクスからクラスタを特定し、SIG意味分析から特定したクラスタの意味づけがなされる。

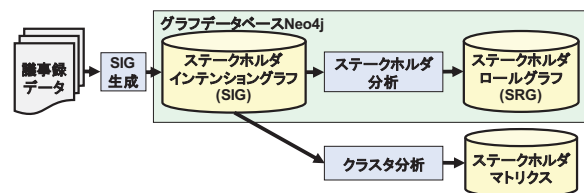


図3 データ間の関係

Figure 3 Relationship between Data

5.2.2 ステークホルダインテンショングラフ (SIG)

ステークホルダインテンショングラフ (SIG) とは、各ステークホルダの発言とその目的語との関係を表したグラフである。本研究におけるグラフ分析対象であり、Cypher による分析を行う。

SIG のメタモデルを図 4 に示す。

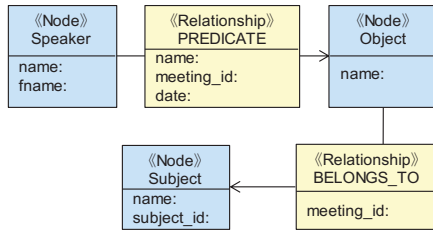


図 4 SIG のメタモデル

Figure 4 Meta Model of Stakeholder Intention Graph

以下にメタモデルを構成する要素について示す。

(1) ノード

青色クラスはノードを表す。各ノードの定義を以下に示す。

a) Speaker

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) Subject

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 (subject_id) を持つ。

(2) リレーションシップ

黄色クラスはリレーションシップを表す。各リレーションシップの定義を以下に示す。

a) PREDICATE

発話者と目的語を結び、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

目的語と議題を結び、目的語がどの議題に属するのかわかっている。プロパティとして、会議番号 (meeting_id) を持つ。

5.2.3 ステークホルダロールグラフ (SRG)

ステークホルダロールグラフ (SRG) とは、各ステークホルダの会議における役割に着目し、ステークホルダ間の役割の関係を表現したグラフである。SRG は、SIG の 2 段階分析の結果として得られるグラフである。

SRG のメタモデルを図 5 に示す。

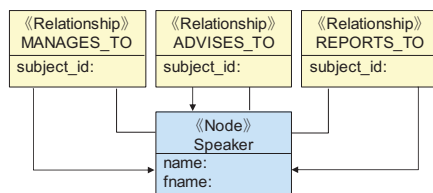


図 5 SRG のメタモデル

Figure 5 Meta Model of Stakeholder Role Graph

以下にメタモデルを構成する要素について示す。

(1) ノード

青色クラスはノードを表す。各ノードの定義を以下に示す。

a) Speaker

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。

(2) リレーションシップ

黄色クラスはリレーションシップを表す。各リレーションシップの定義を以下に示す。

a) MANAGES_TO

発話者と発話者を結び、始点の発話者が終点の発話者を管理する関係を表している。プロパティとして、議事録番号 (subject_id) を持つ。

b) ADVISES_TO

発話者と発話者を結び、始点の発話者が終点の発話者に対して意見や助言を行う関係を表している。プロパティとして、議事録番号 (subject_id) を持つ。

c) REPORTS_TO

発話者と発話者を結び、始点の発話者が終点の発話者に対して報告を行う関係を表している。プロパティとして、議事録番号 (subject_id) を持つ。

5.3 ステークホルダ分析方法

ステークホルダの構造分析方法として 2 段階の分析方法を提案する。

5.3.1 大域的グラフ分析

(1) 分析内容と対象

大域的グラフ分析とは、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である。グラフの大域的構造から議論が活発に行われているステークホルダと目的語を分析する。また、各ノードの次数から議題の中心人物を分析する。

(2) 次数の定義

ノードの次数とは、あるノードと他のノードとのリレーションシップ数であり、有向グラフの場合、ノードの次数は、出次数と入次数が存在する[5]。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評価に有効である。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の評価に有効である。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。

また、Neo4j においてステークホルダノードを固定した場合、目的語ノードはリレーションシップ数が多いステークホルダノードに引かれ、グラフ全体の重心が偏るという特性がある。よって、ノードの次数に基づき、議論が活発に行われているステークホルダと目的語ノードの集合が分析可能である。

5.3.2 局所的グラフ分析

(1) 分析内容と対象

局所的グラフ分析とは、大域的グラフ分析では明らかにならない、特定のステークホルダの集合の意味的構造を明らかに

する分析である。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し、分析をする。

サブグラフとは、大域的グラフ分析で用いたグラフにおける一部のノードとリレーションシップにより形成される部分グラフである。

(2) クリークの定義

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ノードの集合である[11]。本研究では、目的語ノードから直接連結する、測地線距離が 1 のノード(ステークホルダ)の集合を分析対象のクリークとして定義する。

局所的グラフ分析において分析するクリークの範囲の例を図 6 に示す。グラフ中の実線で囲まれた範囲のサブグラフを分析対象のクリークとして扱う。

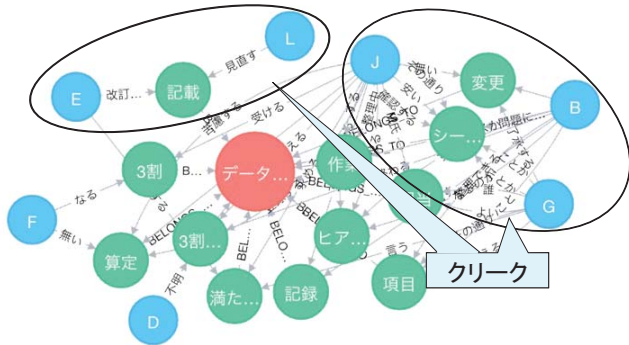


図 6 ステークホルダクリークの範囲例

Figure 6 An Example of Global and Local Graph

(3) ステークホルダクリーク分析

局所的グラフ分析では、抽出したクリークを分析する。抽出したクリークの形状によって、次の 3 つの分析を定義する。

a) 単項ステークホルダ分析

特定のステークホルダ 1 人の関心とその強さを目的語への関係の意味と回数によって分析することで、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する。ステークホルダの共通の目的語に対する発話意図の一致、相違、衝突など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。

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、1 対多のステークホルダ間の相互関係を分析することで、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。

5.3.3 SRG の生成

SIG を分析した結果、明らかになったステークホルダの役割を SIG に追加する。ステークホルダノードとその役割のリレーションシップを抽出し、可視化して、分析結果の SRG とする。

5.3.4 ステークホルダマトリクスの生成と分析

(1) ステークホルダのポジション分析

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定量的分析する。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ため、関与度、影響度を用いる。

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、ステークホルダが議題に関わる程度の指標である。

影響度は、あるステークホルダが他のステークホルダに対して持つ、影響力の程度の指標である。

(2) ステークホルダマトリクスの生成

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。

(3) ステークホルダのクラスタリング

生成したステークホルダマトリクスからクラスタを特定する。SIG のグラフ分析結果より特定したクラスタの意味付けを行う。

6. ステークホルダ分析支援システムのプロトタイプ

議事録データは膨大であり、人手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。したがって、一連の処理を自動化するためのプロトタイプを実装した。

6.1 プロトタイプの実装環境

プロトタイプの実装環境を表 1 に示す。

表 1 プロトタイプの実装環境

Table 1 Implementation's Environment of Prototype

環境	ソフトウェア
OS	macOS Sierra
プロセッサ	2.6 GHz Intel Core i7
メモリ	16 GB
実装言語	Python 3.5.2
形態素解析	MeCab 0.996
係り受け解析	CaboCha 0.69
グラフデータベース	Neo4j 3.0.7

6.2 プロトタイプの実装

図 7 に示すプロトタイプを実装、自動化することで、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。

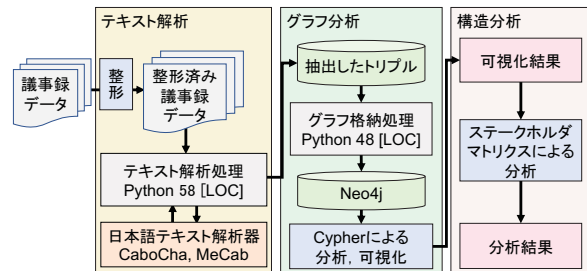


図 7 システム構成図

Figure 7 System Configuration Diagram

(1) テキスト解析プロセス

テキスト解析処理で Python を用いて、58 行で記述した。

(2) グラフ分析プロセス

グラフ格納処理で Python を用いて、48 行で記述した。

7. 実開発への適用

7.1 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。

表 2 に各会議における文章量と実施日を示す。

表 2 各会議における文章量と日付

Table 2 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

7.2 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ため、形態素解析、係り受け解析に適した形式に整形した。

7.3 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、係り受け解析により、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では、発話者名は括弧で囲まれた形で書かれているため、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。

7.4 グラフモデルの生成

Cypher を用いて、図 4 に示した SIG のメタモデルに基づき、データを Neo4j に格納し、グラフモデルを生成する。

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、生成した SIG を可視化した結果を図 8 に示す。

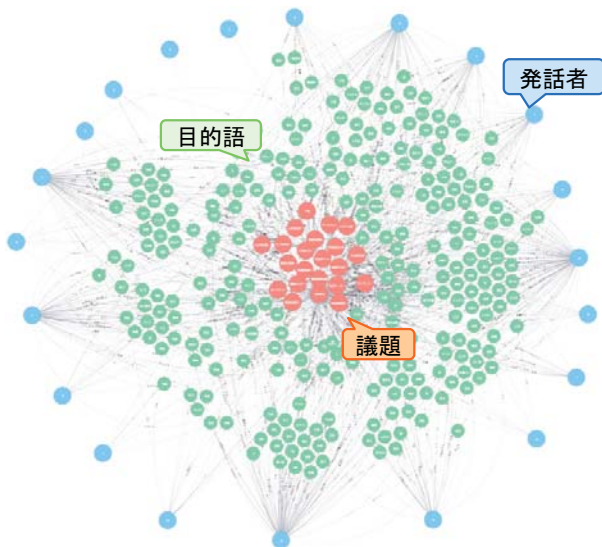


図 8 リレーションシップ追加後のグラフ
 Figure 8 A Graph after Adding Relationship

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。

(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。

格納したノードとリレーションシップの個数を表 3 に示す。

表 3 SIG のノードとリレーションシップの個数

Table 3 Total of Nodes and Relationships of SIG

	名前	個数	合計
ノード	Person	21	317
	Object	274	
	Agenda	22	
リレーションシップ	PREDICATE	754	1198
	BELONGS_TO	444	

7.5 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 8 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、議題毎に関与す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。

7.5.1 分析対象のサブグラフ出力

「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、この議題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。

7.5.2 グラフ分析

本研究で提案したグラフ分析方法を適用する。

(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。

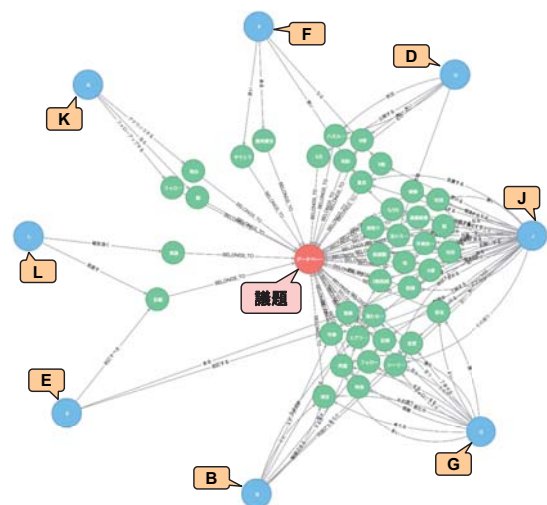


図 9 議題 No.8 データベース登録支援
 Figure 9 Database Registration Support for Agenda No.8

図 9 より、主語ノードを円状に均等に配置しているにも関わらず、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、グラフの重心が偏っていると言える。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。よって、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。

「議題 No.8」における発言数分布を図 10 に示す。グラフの右側に位置する、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。

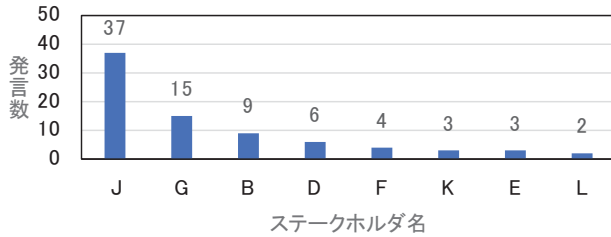


図 10 ステークホルダの発言数分布

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 に抽出したサブグラフを示す。

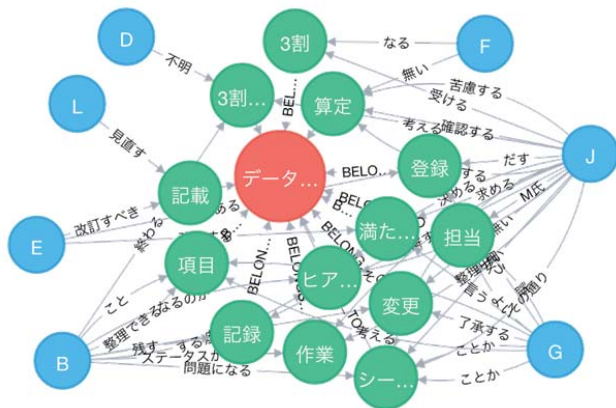


図 11 抽出したサブグラフ

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 9 に含まれているが、図 11 には含まれていない。よって、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への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 の二項関係を図 12 に示す。E の「改定すべき」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。



図 12 E と L の二項関係

Figure 12 Binary Relationship of E and L

ii) D と J、F と J の関係

D と J、F と J の二項関係を図 13 に示す。

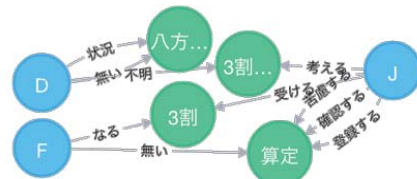


図 13 D と J、F と J の二項関係

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」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、両者の発言ともに 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、こ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。さらに、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。

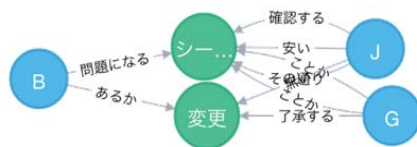


図 14 B、G、J の三項関係

Figure 14 Triadic Interaction of B, G and J

7.6 分析結果の可視化

2 段階分析の結果を踏まえ、図 5 で示したデータモデルに基づき、追加の発話者ステークホルダ間の関係を Neo4j に追加

する。SRG として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。

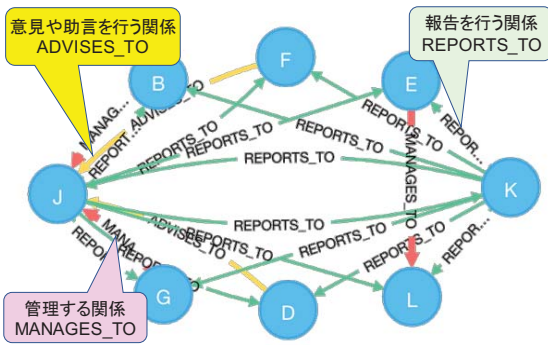


図 15 ステークホルダ間の関係

Figure 15 Relationships between Stakeholders

7.7 ステークホルダマトリクスによる分析

7.7.1 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、J から E, L, K への会議においての一方向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。

7.7.2 発話意図の定義と重み付け

MANAGES_TO 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、ADVISES_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 4 に示す。

表 4 発話意図の重み

Table 4 Weight of Speech Intention

MANAGES_TO	ADVISES_TO
0.8	0.2

7.7.3 関与度と影響度の計算

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 5 に示す。

表 5 各ステークホルダの発言集計結果

Table 5 Sum of Statements of Each Stakeholder

	MANAGES_TO	ADVISES_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 5 において、MANAGES_TO, ADVISES_TO の列で、各ステークホルダの各々の発言数を示す。また、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。

7.7.4 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を生成し、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。この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。この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職は、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 SME (Subject Matter Expert) [4] 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関する報告を行うステークホルダが含まれる。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、関与度と影響度から報告者クラスタに分類されるステークホルダは、議題に深く関与して上位の立場のステークホルダに報告を行うステークホルダであると推定される。

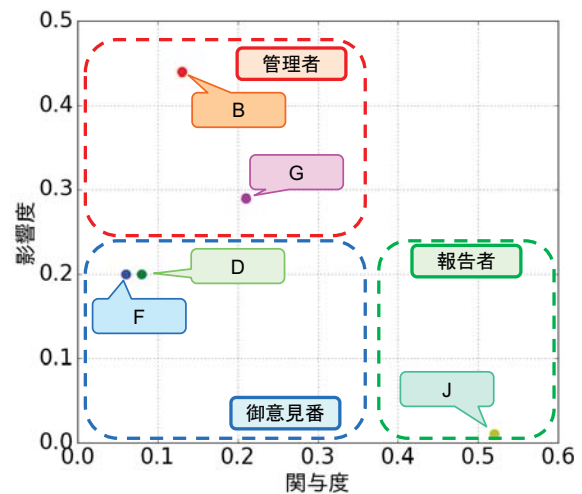


図 16 ステークホルダマトリクス

Figure 16 Stakeholder Matrix

8. 評価

8.1 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。

8.2 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

ステークホルダ分析方法として、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。

(1) 大域的グラフ分析の評価

大域的グラフ分析では、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。また、グラフの形状から、どの話題が注目されているかの話題ステークホルダの分析も可能である。

(2) 局所的グラフ分析の評価

局所的グラフ分析では、SIGから共通の目的語を持つステークホルダの集合をクリークとして抽出し、意味分析を行うことで各ステークホルダの役割を分析可能であることを示した。これにより、SRGを生成し、ステークホルダの構造化が可能になった。

(3) ステークホルダマトリクスの生成と分析の評価

発話データのステークホルダの発話意図に基づき、影響度と関与度を定量的に評価した。この評価結果からステークホルダマトリクスが自動生成可能となり、データに基づいた合理的なステークホルダ分析が可能である。

8.3 実データへの適用による提案方法の妥当性評価

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のデータ量が増加すると考えられるが、セマンティックグラフモデル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。

9. 考察

9.1 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[3]では、分析が困難である複雑な発話データからステークホルダ間の構造を定量的に分析可能である。

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データに基づいた要求獲得が可能となり、要求獲得の効率化と要求仕様の品質向上が期待できる。

9.2 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[2]が提案されているが、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方法を提案した。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、ステークホルダの役割が特定できる。

9.3 従来のRDBを用いた分析との比較

システム開発において、会議の進行とともに議題毎に関連するステークホルダの変更や発話意図の変化がある。よって、その場合、従来のRDBではデータモデルの再定義が必要になる。しかし、セマンティックグラフモデル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分析が行える拡張性がある。さらに、同一モデル上でステークホルダの発話意図の変化とそれに伴う議事の変化を分析できる。

10. 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の特性を利用した分析を検討する。

(2) セマンティックグラフモデルの適用範囲の拡大

提案方法の適用範囲を他の発話データや議事録データに拡大し、提案方法の妥当性を確認する。

(3) 機械学習技術の導入

機械学習技術による大量の発話データからのステークホルダの発話意図の分析方法を検討する。

11. 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構造の意味分析方法の提案をした。

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RDBでは表現できないステークホルダ間の意味関係に着目した動的な構造分析が可能となった。また、議事録データからステークホルダマトリクスを自動生成可能となり、クラスタリングした結果からステークホルダの役割が特定可能である。したがって、データに基づく合理的で高品質な要求獲得につながる支援が可能となる。

謝辞 議事録データをご提供頂いた伊藤忠テクノソリューションズ株式会社の野村典文氏と関係各位に感謝する。

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] 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.
- [4] JISA REBOK企画WG(編), 要求工学知識体系, 近代科学社, 2011.
- [5] 小林 みどり, あたらしいグラフ理論入門, 牧野書店, 2013.
- [6] MeCab, Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/>.
- [7] 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, 電子情報通信学会, SIGKBSE, No. 93-41, Jan. 1994, pp. 9-16.
- [8] Neo Technology, neo4j, 2016, <http://neo4j.com/>.
- [9] I. Robinson, et al., Graph Databases, 2nd ed., 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, 新曜社, 2001.