

データマイニングによる RFP 内部トレーサビリティ分析方法の提案と評価

M2016SE001 朝岡 稔

指導教員 青山 幹雄

1. はじめに

提案依頼書(RFP: Request For Proposal)は、ソフトウェア開発を委託するに当たり、委託元(企業、公共事業者など)のユーザが委託先のベンダに対して開発に関する具体的な技術提案(技術仕様、技術提案書の作成)を依頼する文書である。RFP に記載された内容はソフトウェア開発のベースとなりその品質が後工程の成否を大きく左右する。

2. 研究の背景

RFP はユーザ視点とシステム化視点への対応が必要である。しかし、提示された RFP に基づきシステム化視点で企画すると、RFP のあいまいさが発見される場合がある。そのため繰り返し要求との関連を評価する反復型のプロセスが求められる[7]。このプロセスの詳細を図1に示す。

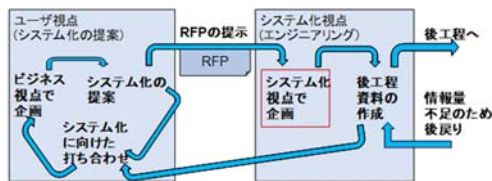


図1 反復的プロセス

3. 関連研究

3.1. 要求の獲得プロセス

現状システムから要求を獲得するために以下の2つのモデルを構築した上で要求の妥当性を確認しシステムに割り当てるプロセスが提案されている[4]。

- (1) 現状のシステムモデル
- (2) 将来のシステムモデル

3.2. 非機能要求グレード

非機能要求グレードは非機能要求の項目を階層的にグループ化し、段階的な評価基準として提案されている[2]。

3.3. 政府情報システムの標準ガイドライン

政府による情報システム開発の標準ガイドラインとして「政府情報システムの整備及び管理に関する標準ガイドライン」(以下政府ガイドラインと略記)が規定されている[5]。

3.4. 非機能要求に着目した RFP 評価

ユーザが運用と保守に関する非機能要求の記述の有無および明確さによって RFP を定量的に評価する方法が提案されている[8]。

4. アプローチ

RFP の内部トレーサビリティを発見するためのデータマイニングの手法を提案する。

図2にデータマイニングのアプローチを示す。懸念になることはRFPの作成過程上の業務見直しで獲得した業務要求が要求定義に反映されないことである。要因として業務要求に結び付く要求が発見できないことが指摘できる。

解決策としてクラスタリングを用いて非機能要求を基準に分類することで、ユーザ視点とシステム化視点の間で共通の非機能要求が発見する方法を提案する。発見されたクラスタ単位でトレーサビリティの水平リンクを構築することで非機能要求の管理が可能になる。

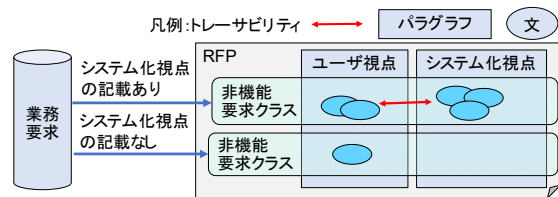


図2 トレーサビリティのイメージ

5. RFP 内部トレーサビリティ分析方法

5.1. RFP の文書構造の解析

本稿では多様な要求のなかで、業務要求とシステムに割り当てられる非機能要求に焦点をあてる。

RFP 見本[3]をベースとすると、親パラグラフ、子パラグラフ、文、非機能要求の順に親子関係をもつ木構造で構成される。適用した結果を図3に示す。

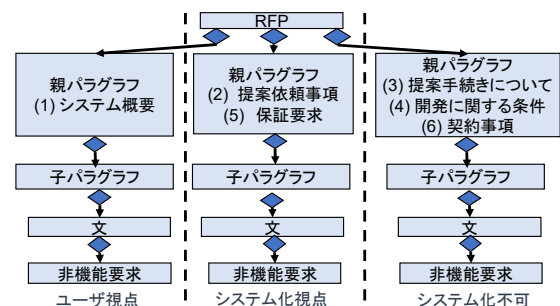


図3 開発 RFP 見本の分類結果

以下の分類が可能である。

- (1) ユーザ視点: 業務要求について記載されたパラグラフ
- (2) システム化視点: システムの各種要求について記載されたパラグラフ
- (3) システム化不可: システムと無関係な事柄について記載されたパラグラフ

5.2. トレーサビリティ構築のプロセス

評価プロセスは一般の知識発見の手法に従い、以下の手順で行う。詳細は図4を参照。

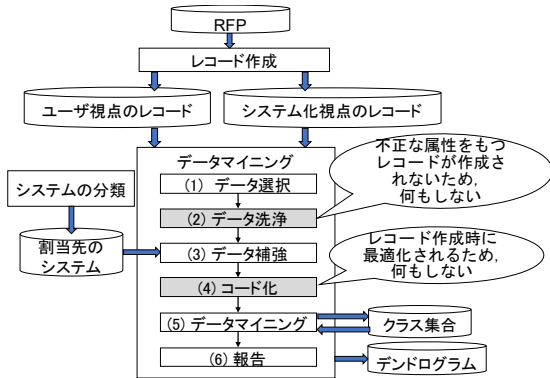


図4 分析プロセス

- (1) データ選択: どのデータを評価するかを決める。
- (2) データ洗浄: 不正なデータまたは無意味なデータを除去する。
- (3) データ補強: 外部データを利用して各データを適切な形に強化する。
- (4) コード化: 「データ」を「データマイニング」を行うに適した形に変換する。
- (5) データマイニング: 実際の知識を発見する。
- (6) 報告: データマイニングの結果をまとめる。

5.3. 入力レコードの形成

本稿で提案するデータマイニングに使用するレコードの構造を表1に示す。

表1 入力レコードの構造

No	項目
1	評価視点
2~88	非機能要求(小項目)
89~130	非機能要求(中項目)
131	クラスラベル

5.4. RFPの持つシステムを構成するサブシステムの分類

システム視点のみで判断できるシステムはユーザ視点と共有できないため割当先のサブシステムから除外する。よって評価される文はユーザ視点と共有可能なサブシステムに対して言及している文に特定される。

5.5. 非機能要求の抽出

政府ガイドラインに沿った非機能要求に絞って非機能要求の抽出を行う。配点は非機能要求を具体的に記載して

いるかどうかを観点に行う。そのための評価方法を以下に列挙する。

(1) 記載内容の評価

システム基盤について考慮されず非機能要求が達成された前提で記載されることが考えられる。業務要求を達成するための前提条件の非機能要求を評価することで手段として検討されるユーザ視点の非機能要求を獲得が可能である。

機能要求を達成するための前提条件の非機能要求も同様に評価することでシステム化視点の非機能要求の獲得が可能である。詳細は図5を参照。

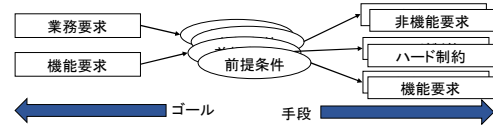


図5 暗黙的な非機能要求の図示

(2) クラス間の距離の抽出

ユークリッド距離による算出に加えて非機能要求の中項目と以下の項目間の算出ロジックを変更する。

- 1) 直下の小項目
- 2) 異なる中項目直下の小項目

以下に階層構造を考慮した距離の算出を図6に示す。階層を考慮することで「1)直下の小項目」間の距離が「2)異なる中項目直下の小項目」よりも近くなる。

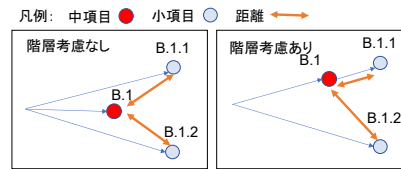


図6 階層構造と距離

6. 適用

6.1. 検証対象に関する仮説

業務要求の内容が後工程に展開されることから、適用方法を式1で定義した正確度で評価する。

$$\text{正確度} = A \div (A + B) \quad (1)$$

ただし、AとBは下記で定義する。

A: ユーザ視点の要素とシステム化視点の要素を含むクラスの文の数

B: ユーザ視点の要素を含むがシステム化視点の要素を含まないクラスの文の数

6.2. 内部トレーサビリティの適用

以下の3つのRFPに対して適用を実施した。

- (1) 葛飾区行政情報配信等アプリ構築委託提案依頼書 (RFP 1)[6]
- (2) 横須賀市仮想デスクトップ基盤構築業務提案依頼書

(RFP 2)[9]

(3) 企業情報管理システム開発委託提案依頼書(RFP 3) [1]

6.3. 基礎統計

6.3.1. RFP 文書の量

各 RFP の文の数を文単位で示す。適用対象および適用対象外の文について表 2 にまとめた。

表 2 適用対象の文書の規模

項目	RFP 1	RFP 2	RFP 3
適用対象の文	187	72	128
適用対象外の文	63	74	70
文書規模(文)	250	146	198

6.3.2. 構造

パラグラフを分類し色分けした結果を図 7 に示す。

- (1) 赤: ユーザ視点のパラグラフ
- (2) 青: システム化視点のパラグラフ
- (3) 灰色: システム化不可のパラグラフ
- (4) 白: 子を持つパラグラフ

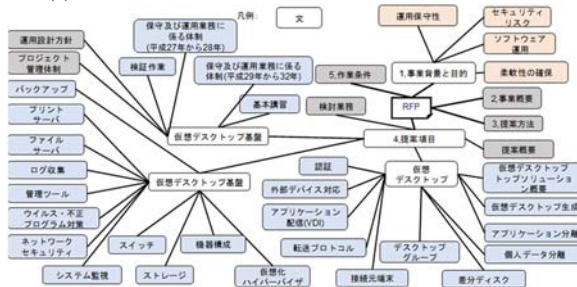


図 7 RFP2 の文書構造グラフ

6.4. 非機能要求の抽出

6.4.1. 分析の精度

非機能要求の分析の精度の検証のためユーザ視点のレコードの統計を表 3 にまとめた。抽出率を式 2 で定義する。

$$\text{抽出率(\%)} = P \div (P + Q) \quad (2)$$

ただし、P と Q は下記で定義する。

P: 非機能要求の抽出が可能なた文

Q: 非機能要求の抽出が不可能な文

表 3 ユーザ視点の非機能要件の抽出

項目	RFP 1	RFP 2	RFP 3
P	4	12	2
Q	3	6	11
P+Q	7	18	13
抽出率	57.1	66.7	15.4

6.4.2. 抽出結果

ユーザ視点とシステム化視点で抽出した非機能要求の項目と数について図 8~10 にまとめた。

図に示した割合の式 3 を以下に示す。

$$\text{評価視点毎の非機能要求率} = S \div T \quad (3)$$

ただし、S と T は下記で定義する。

S: 各視点(ユーザ視点またはシステム化視点またはユーザ視点とシステム化視点両方)の非機能要求の数

T: RFP で抽出した非機能要求の数

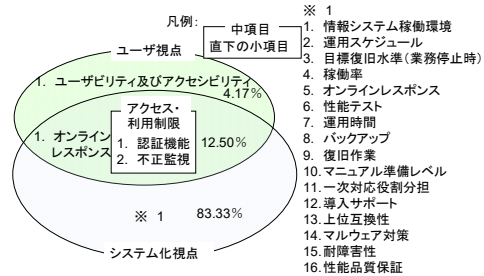


図 8 RFP1 を評価視点で分類した Venn 図

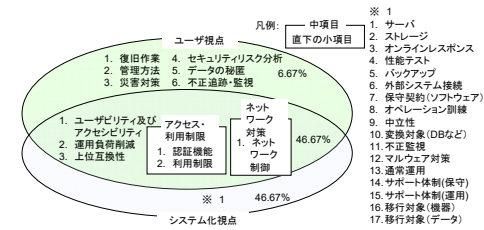


図 9 RFP2 を評価視点で分類した Venn 図

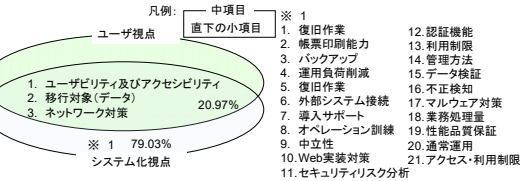


図 10 RFP3 を評価視点で分類した Venn 図

6.5. トレーサビリティの構築

分類結果を図 11~13 に示す。

(1) RFP1 の分類結果

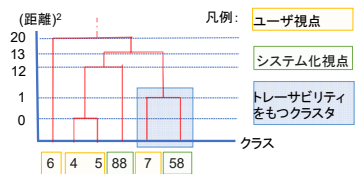


図 11 RFP1 のデンドログラム

(2) RFP2 の分類結果

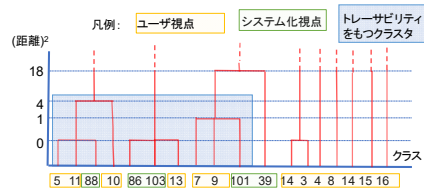


図 12 RFP2 のデンドログラム

(3) RFP3 の分類結果

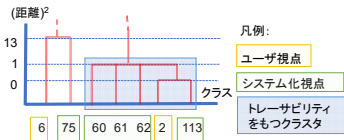


図 13 RFP3 のデンドログラム

6.6. 分類結果

ユーザ視点で構築した非機能要求の構築の正確度を検証するため、分類結果を表 4 にまとめた。

表 4 トレーサビリティ統計

項目	RFP 1	RFP 2	RFP 3
トレーサビリティが抽出可能なユーザ視点の文の数	1	6	1
トレーサビリティが抽出不可能なユーザ視点の文の数	3	6	1
式 2 で算出した正確度	25.0	50.0	50.0

7. 評価

適用結果を非機能要求の抽出とトレーサビリティの取得の 2 つの観点で評価する。

7.1. 非機能要求の抽出

ユーザ視点の非機能要求とシステム化視点の非機能要求を比較することで、ユーザ視点の非機能要求をシステム化視点へ展開されているかどうかを確認することが可能であることが明らかになった。

7.2. トレーサビリティ分析の効果

システム化視点の非機能要求はユーザ視点の非機能要求から展開された要求であると仮定したが、正確度が低いことが明らかになった。

7.3. 提案分析方法の評価

非機能要求の抽出が可能であるが、提案した方法ではトレーサビリティの取得の精度が低いことが明らかになった。その要因を 2 つ挙げる。

- (1) ユーザ視点の非機能要求の分析の過程で 1 つの文から複数の非機能要求が抽出されるため。
- (2) RFP にユーザ視点の非機能要求を用いてシステム化視点で分析した結果を記載していないため。

8. 考察

先行研究と本稿との比較を表 5 に示す。

先行研究[8]では運用と保守を対象に 55 項目の非機能要求の記述の明確さを 5 段階で人手により評価している。それに対して、本稿では非機能要求グレードで定義されている非機能要求の項目を政府ガイドラインに基づき絞り込んで 129 項目の非機能要求を対象にしている。また、本稿ではデータマイニングによって分類を自動化することによって、合理的な評価を可能にしている。

表 5 本稿と先行研究の比較

項目	目的	評価対象	評価観点
先行研究	非機能要求の記述の指標の取得	ガイドライン上の非機能要求	非機能要求の明確性
本稿	スコープ間のトレーサビリティの取得	ユーザ視点で抽出された非機能要求	必要性および実現手段

本稿は、ユーザ視点とシステム化視点の間の非機能要求のトレーサビリティを特定することで非機能要求の必要性と実現手段を評価することを可能としたことに意義がある。

9. 今後の課題

今後、下記の課題を検討する必要がある。

- (1) クラスタリングの精度を上げるため、クラスタリング手法および入力レコードの属性の検討が必要である。
- (2) 本稿では内部トレーサビリティを扱ったが外部トレーサビリティに適用する手法を検討する必要がある。

10. まとめ

現行の RFP の分析において、クラスタリングを用いて文書構造を分析する方法を提案した。このため、ユーザ視点とシステム化視点の間の非機能要求のトレーサビリティを特定することで非機能要求の必要性と実現手段を評価する方法を提案した。非機能要求の項目として政府の発行したガイドラインをベースにしてデータマイニングによる分類を行う基準を定義した。

Web に公開されている 3 件の公共情報サービスシステムの RFP に対して提案方法を適用し、非提案方法の有効性、妥当性を評価した。

参考文献

- [1] 台東区産業振興事業団, 企業情報管理システム開発委託提案依頼書, 2017.
- [2] IPA, 非機能要求グレード利用ガイド[解説編], 2010.
- [3] IT コーディネータ協会, 開発委託用 RFP 見本, 2014.
- [4] 飯村 結香子 ほか, REBOK に基づく要求分析実践ガイド, 近代科学社, 2015.
- [5] 各府省情報化総括責任者(CIO)連絡会議, 政府情報システムの整備及び管理に関するガイドライン, 2014.
- [6] 葛飾区, 行政情報配信等アプリ構築委託提案依頼書, 2016.
- [7] 宮脇 亨, マイルストーン計画と断面評価による段階的な要件定義の進め方, UNISYS Technology Review, No. 122, Dec. 2014, pp. 13-24.
- [8] 齊藤 康廣, ほか, 非機能要件に着目した RFP 評価, SEC Journal, Vol. 10, No. 3, Sep. 2014, pp. 30-37.
- [9] 横須賀市, 仮想デスクトップ基盤構築業務提案依頼書, 2015.