Vol.2019-CSEC-86 No.66
Vol.2019-SPT-34 No.66
2019/7/24

情報セキュリティに関連するガイドラインの Doc2Vec を用いた文書内容の可視化手法の提案とその評価

尾崎敏司[†1]

概要:2012 年に独立法人情報処理推進機構により提示された「情報キュリティ人材の育成に関する基礎調査」と 2014 年のその追加調査によると、約 8.1 万人の情報セキュリティの人材不足が指摘されており、現在もその育成は課題となり続けている。自己学習の起点となると考えられるガイドラインは多く公開されているが、これらのガイドラインがセキュリティ業務のどの部分に該当するのか初学者が把握するのは難しい。そこで、本研究では米国国立標準技術研究所の公開している Cybersecurity Framework1.1 をもとに、Doc2Vec を用いてガイドラインの文書内容を体系的に提示する手法の提案を行い、質的コーディングを実施した結果と比較することでその評価を行った。

Proposal to visualize a content of information security guideline based on Cybersecurity Framework with Doc2Vec approach

SATOSHI OZAKI[†1]

1. はじめに

2012 年に独立法人情報処理推進機構（IPA）により提示された「情報セキュリティ人材の育成に関する基礎調査」[1]と 2014 年に行われた追加分析[2]によると，約 8.1 万人の情報セキュリティの人材不足が指摘されており，現在もその育成は課題となり続けている．また，内閣サイバーセキュリティセンターでは，サイバーセキュリティ人材の育成に関する施策連携ワーキンググループが結成されており，2018 年に「サイバーセキュリティ人材の育成に関する施策間連携ワーキンググループ 報告書」[3] が作成されている．この報告書では，セキュリティの専門家であるスペシャリストと，一般的な社内の IT オペレーションを実施しているゼネラリストの間に，エキスパートと呼ばれる「自社事業とセキュリティ活動をよく知り，現場と経営をつなぐ人材」の必要性を指摘しており，引き続き企業における人材育成が求められていることが伺える．

前述の「情報セキュリティ人材の育成に関する基礎調査」の追加分析によると，約8.1万人の情報セキュリティの人材不足のうち，現在セキュリティ人材を保持していない企業において新たに必要とされる人数は6.1万人と推計されている．同時期に情報セキュリティ大学院大学により行われた「情報セキュリティ事故対応に関わるアンケート調査」[4] の結果においても，無回答層を含めた場合，中小企業における約75%がセキュリティ担当者を保持していない可能性が示唆されており，担当者をおいている場合でも約41％が兼任の担当者1名のみの状態であった．トレンドマイクロ株式会社が2018年9月に発行した「法人組織におけるセキュリティ実態調査2017年版」[5] においては，従業員規模とセキュリティ対策の包括度に相関関係があることが指摘されており，特に，中小企業において引き続き限られた人材・資源の中でセキュリティ対策を実施していくことが必要になると考えられる．

2. 関連研究

セキュリティ人材の教育手法の研究としては，近年では，CTF (Capture the Flag) によるアプローチに関した研究が多くみられる[6][7][8]．また，攻撃手法の学習上課題となることが多い演習環境に注目したものも多い [9][10].

これらの教育手法は，技術的な側面での学習補助として有用であると考えられる．しかし，エキスパートつまり「自社事業とセキュリティ活動をよく知り，現場と経営をつなぐ人材」という観点では，技術に限らない広い視点での学習活動が求められている．これを実現するためには，学習者の主体的な学習を補助するアプローチが必要になると考

†1 筑波大学
University of Tsukuba

Vol.2019-CSEC-86 No.66
Vol.2019-SPT-34 No.66
2019/7/24

えられる.

エキスパートの育成を目的とした研究では，孫らによる大学・大学院のカリキュラムに対する研究が挙げられる[11]. 孫らは，アメリカ国立標準技術研究所（NIST）の下に設置されている NICE（The National Initiative For Cybersecurity Education）が定義した Cybersecurity Workforce Framework[12] に基づいて大学と大学院におけるセキュリティ教育課程カリキュラムの分析を行った. この研究では，Cybersecurity Workforce Framework の 783 個の技術能力項目を 62 種類の項目に集約している. この 62 の項目について，a) Cybersecurity Workforce Framework 内の単語出現数でつけた 62 項目の順位と，b) 大学・大学院カリキュラムと 62 項目の対応付けを行い科目の数でつけた順位の二つの順位を作成し a) と b) の間の Spearman 順位相関係数を用いて，大学・大学院カリキュラムの妥当性を検証している.

この研究は，大学・大学院教育を対象に，カリキュラム開発の要求分析を目的として， Cybersecurity Workforce Framework に基づいた項目を単語の出現数や科目数で順位づけて比較を行い, 教育課程全体の妥当性を評価している.

尾崎[13] は，主に社会人の学習者に利用されるセキュリティ関連のガイドラインを対象に，学習者の主体的な学習の補助や実務実施の補助を目的として，NIST の発行している Cybersecurity Framework 1.1[14] に基づいて tf-idf（Term Frequency-Inverse Document Frequency）を用いて分析行い，文書毎のその文書内容の可視化を行っている.

しかしながら，tf-idf は文書中の単語の語数や出現頻度に基づいた分析手法であり，語順などの要素は解析結果に影響を与えない. そこで，語順などの情報が含まれるベクトル表現である Doc2Vec の PV-DM（Paragraph Vector-Distributed Memory）[15] で分析を行うことを検討する.

2.1 自己調整学習に関する研究

自己調整学習とは，1990 年代からアメリカの教育心理者 Barry Zimmerman らが中心となって提案している教育心理学の理論体系で，学習者の主体的な学習方略を重要視している. 学習方略とは，学習に取り組む戦略のことで，自己調整学習では，学習方略を大きく，認知的方略（例：関連付けて覚える），メタ認知的方略（例：勉強時間と学習範囲を記録する），動機付け方略（例：学習の目的を書き出す）に分けている[16].

このうち認知的方略に含まれる体制化方略と図示化方略は，国内では，松村による英語の現在完了形の学習における実験で、学習効果の向上に寄与することが確認されている[17]. 体制化方略とは，「何らかの理論や枠組みによって学習要素を相互に関連付けて整理する方法」であり，図示化方略とは，文字通り「図示により整理」を行う方法である.

本研究では，他分野ではあるが先行研究で効果が確認されている体制化方略につながる形での情報提示を検討した.

3. 本研究の目的

本研究では，セキュリティ関連のガイドラインに関して，学習者の体制化方略を補助することを目的として，Cybersecurity Framework 1.1 に基づき，Doc2Vec を用いて可視化する方法について検討する，

4. 提案手法

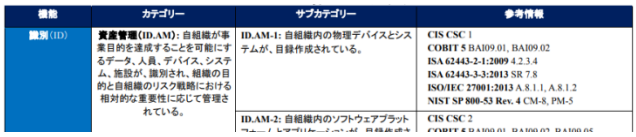
尾崎の先行研究では，全体像を意識した体系的な内容の提示を行うため， Cybersecurity Framework 1.1 の「フレームワークコア」と呼ばれるモデルを基に，tf-idf を用いて文書内容の提示を行うことを検討している.

しかしながら，tf-idf は文書中の単語の語数や出現頻度に基づいた分析手法であり，語順などの要素は解析結果に影響を与えない. 本研究では，語順などの情報が含まれるベクトル表現である Doc2Vec の PV-DM を用いて分析を行うことを検討する.

4.1 Cybersecurity Framework

このフレームワークは，重要インフラストラクチャにおけるセキュリティ対策向けに作成されており，現在，産業界で効力を発揮している標準，ガイドライン，およびベストプラクティスを集約することで，多様なサイバーセキュリティアプローチを体系化・構造化し示している. 解析の対象とする文書は日本語であるため，IPA が発行している本文書を翻訳した「重要インフラのサイバーセキュリティを向上させるためのフレームワーク 1.1[18]」を用いた.

このフレームワークで提示されている「フレームワークコア」は，機能，カテゴリ，サブカテゴリ，参考情報の四つで構成されており，先行研究の提案手法では主に「フレームワークコア」の機能とカテゴリが利用されている. 機能は，基本的なサイバーセキュリティ対策の最も上位の構成要素として「特定」,「防御」,「検知」,「対応」,「復旧」の 5 つが定義されており，カテゴリは各機能をさらに効果毎に分類したものである.



機能	カテゴリー	サブカテゴリー	参考情報
識別（ID）	資産管理（ID.AM）: 自組織が事業目的を達成することを可能にするデータ、人員、デバイス、システム、施設が、識別され、組織の目的と自組織のリスク戦略における相対的な重要性に応じて管理されている。	ID.AM-1: 自組織内の物理デバイスとシステムが、目録作成されている。	CIS CSC 1 COBIT 5 BAI09.01, BAI09.02 ISA 62443-2-1:2009 4.2.3.4 ISA 62443-3-3:2013 SR 7.8 ISO/IEC 27001:2013 A.8.1.1, A.8.1.2 NIST SP 800-53 Rev. 4 CM-8, PM-5
		ID.AM-2: 自組織内のソフトウェアプラット	CIS CSC 2

図 1 重要インフラのサイバーセキュリティを改善するためのフレームワーク 1.1 版の「表 2 フレームワークコア」より部分的に引用

Vol.2019-CSEC-86 No.66
Vol.2019-SPT-34 No.66
2019/7/24

4.2 中小企業の情報セキュリティ対策ガイドライン[19]

このガイドラインは，中小企業の IT 利用の活用が進む中で中小企業がセキュリティ対策に取り組むための指針として 2009 年に作成され，2017 年に法改正等最新の情報を基に改定されたものである．このガイドラインには，チェックリストなどが同梱されており，学習目的のみだけでなく実際にガイドラインに基づいた運用を行えるよう工夫がされている．特に問題を抱えていると思われる中小企業のセキュリティ担当者が最初にふれるドキュメントであろうと考えられるため，今回の解析・評価の対象とした．

機能の識別子	機能	カテゴリーの識別子	カテゴリー
ID	識別	ID.AM	資産管理
		ID.BE	ビジネス環境
		ID.GV	ガバナンス
		ID.RA	リスクアセスメント
		ID.RM	リスクマネジメント戦略
		ID.SC	サプライチェーンリスクマネジメント
PR	防御	PR.AC	アイデンティティ管理とアクセス制御
		PR.AT	意識向上およびトレーニング
		PR.DS	データセキュリティ
		PR.IP	情報を保護するためのプロセスおよび手順
		PR.MA	保守
		PR.PT	保護技術
DE	検知	DE.AE	異常とイベント
		DE.CM	セキュリティの継続的なモニタリング
		DE.DP	検知プロセス
RS	対応	RS.RP	対応計画の作成
		RS.CO	コミュニケーション
		RS.AN	分析
		RS.MI	低減
		RS.IM	改善
RC	復旧	RC.RP	復旧計画の作成
		RC.IM	改善
		RC.CO	コミュニケーション

図 2 重要インフラのサイバーセキュリティを改善するためのフレームワーク 1.1 版の「表 1 機能とカテゴリーの識別子」を引用

4.3 Doc2Vec を用いた提案手法

各カテゴリに関連した記述が解析対象の文書内にどの程度存在するのか可視化するために，Doc2Vec によるベクトル表記を利用し，各カテゴリと解析対象の文書の間のコサイン類似度を算出する（図 3）．具体的には，下記の手順を用いて計算した．

1.「重要インフラのサイバーセキュリティを向上させるためのフレームワーク 1.1」の文章を，表題，句点で分割し，各部分を学習用の文書としてモデルを作成した．

2.「重要インフラのサイバーセキュリティを向上させるためのフレームワーク 1.1」内の各カテゴリに関連した記述を集めたものをカテゴリ毎の文書とした．

3. 学習済みのモデルを基に解析対象の文書とカテゴリ毎の文書の間のコサイン類似度を計算した．

この際，2 通りの方法で計算を実施した．

a. 解析対象の文書全体を 1 文書として，コサイン類似度を計算する．

b. 解析対象の文書の 1 文ずつを 1 文書としてみなして，

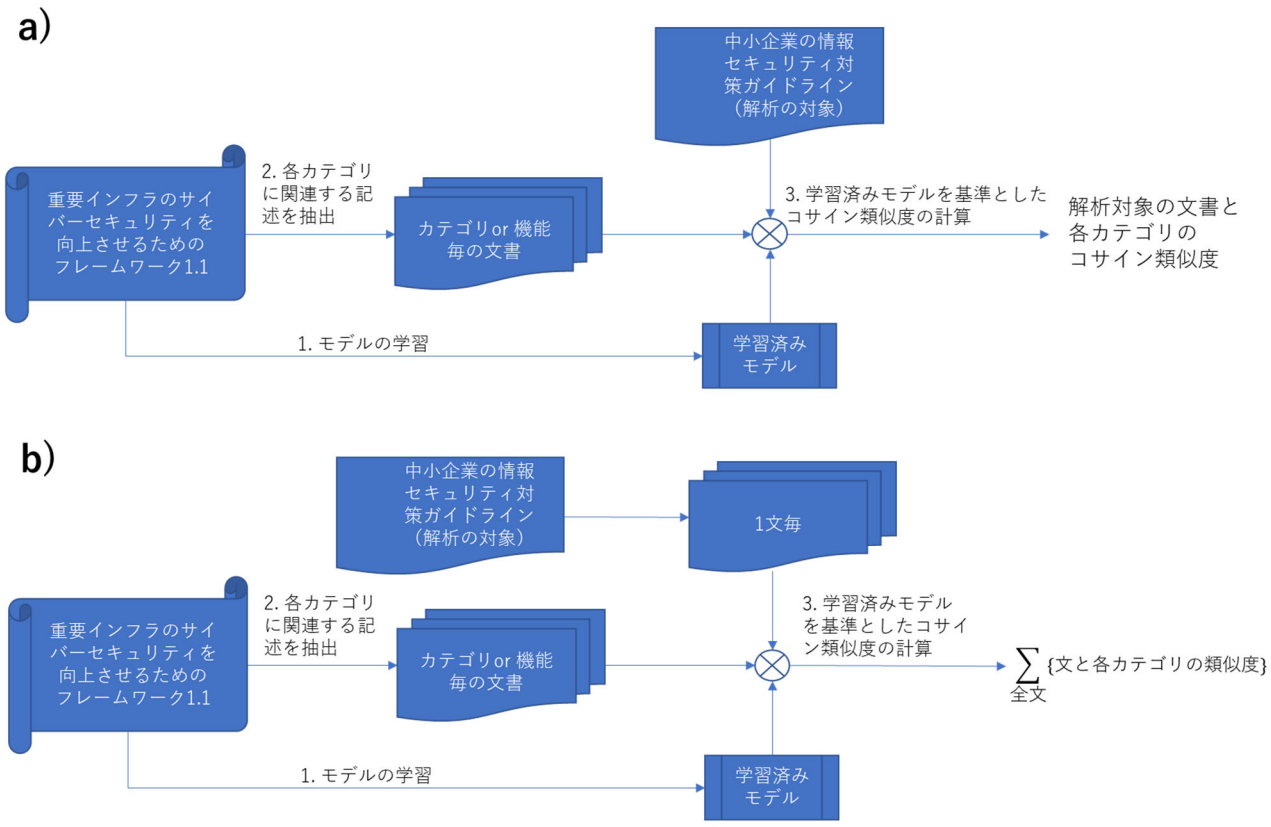


図 3 Doc2Vec による提案手法の手順

Vol.2019-CSEC-86 No.66
Vol.2019-SPT-34 No.66
2019/7/24

1 文ずつコサイン類似度を計算し，その総和をとる．

また，本研究では，カテゴリで実施した場合と同様に，機能ごとに記述を抽出して，同様にコサイン類似度の計算を実施した．

分かち書きには Mecab[20] を，Doc2Vec のモデル作成と文書間のコサイン類似度の計算には gensim [21] のライブラリを利用している．モデルの学習は，alpha=0.0015（学習率），min_count=1（単語の最低出現回数），sample=6（頻出単語を無視する閾値），vector_size=300（ベクトル化した際の次元数）で実施した．

4.4 質的コーディングによる評価用のデータについて

質的コーディングの結果は，尾崎の先行研究のものを用いるが，その手法について記載を行う．

提案手法の精度を測定するためには，「解析対象の文書の内容を人がどのように理解しているか」を定量的に表すことが必要とされる．そこで，質的コーディングの方法を用いて文書の内容を人の手で分析・定量化することで，提案手法の評価に用いるデータとした．

質的コーディングとは，質的データ解析の方法の一つで，文書に表れる表現に対してコード（符号）を割り当てることで，文書の内容について整理を行う手法である．事前にコーディングに用いる語群（コード群）を定義するテンプレートコーディングと，繰り返し文書を読みながら都度コードを作成していくオープンコーディングに大別される．

提案手法の評価を行うには，提案手法と比較可能な形式で文書の質的解析を行う必要がある．そこで，Cybersecurity Framework 1.1 のフレームワークコアのサブカテゴリをコード群として，解析対象にした 4 つの文書に対してテンプレートコーディングを実施した．この結果を提案手法の結果と比較することで，提案手法の評価が可能になる．

コーディングを実施する際には，

a) 原則，1 センテンスごとに評価を行う．「用語の説明＋用語を用いた文」，「説明＋補足事項」などの 2 つ以上のセンテンスで一つの意味を

表 1 提案手法によるスコアと質的コーディングの結果

	提案手法による解析							質的コーディング		
	機能		機能ごとカテゴリ平均		カテゴリ					
機能	b. 1行毎の結果の総和	a. 1文書	b. 1行毎の結果の総和	a. 1文書	b.1行毎の結果の総和	a. 1文書	カテゴリ	スコア	正規化スコア	機能ごとの平均値
Identify（特定）	24.5395	0.9164	17.3580	0.6787	21.3960	0.8382	資産管理	38	0.6129	28.3333
					12.0298	0.5619	ビジネス環境	8	0.1290	
					10.9967	0.3266	ガバナンス	52	0.8387	
					26.2314	0.9548	リスクアセスメント	41	0.6613	
					17.8758	0.7659	リスク管理戦略	5	0.0806	
					15.6183	0.6247	サプライチェーンリスク	26	0.4194	
Protection（防御）	27.0271	0.9699	20.0615	0.7359	18.0991	0.6276	アクセス制御	8	0.1290	20.1667
					16.6970	0.5607	意識向上およびトレーニング	39	0.6290	
					17.5934	0.6619	データセキュリティ	10	0.1613	
					26.2795	0.9676	情報を保護するためのプロセスおよび手順	62	1.0000	
					18.6654	0.7318	保守	0	0.0000	
					23.0348	0.8659	保護技術	2	0.0323	
Detection（検知）	26.4543	0.9665	18.0250	0.6844	25.2195	0.9205	異常とイベント	0	0.0000	2.3333
					10.8767	0.4438	セキュリティの継続的なモニタリング	7	0.1129	
					17.9788	0.6890	検知プロセス	0	0.0000	
Response（対応）	26.5830	0.9661	11.5023	0.5171	12.3644	0.5393	分析	2	0.0323	6.8000
					2.9187	0.1953	コミュニケーション	17	0.2742	
					-1.5881	0.1245	改善	10	0.1613	
					24.6850	0.9422	低減	2	0.0323	
					19.1312	0.7840	対応計画	3	0.0484	
Recovery（復旧）	20.7813	0.5969	9.4584	0.3508	11.7645	0.3666	改善	9	0.1452	3.6667
					11.8964	0.4771	コミュニケーション	0	0.0000	
					4.7144	0.2087	復旧計画	2	0.0323	

Vol.2019-CSEC-86 No.66
Vol.2019-SPT-34 No.66
2019/7/24

成していると考えられた部分には，そのまとまりでの評価を実施している．

b) 複数のサブカテゴリに該当すると考えられた場合には，複数のコードを割り振る．

c) 図表など，画像として添付されている項目はコーディングの対象に含めない．

d) コード群に適切なコードが存在しないと思われる場合には，その文に対してコードの割り振りは実施しない．

こととした．

例えば，「中小企業の情報セキュリティ対策ガイドライン」の中にある「パソコンにはウイルス対策ソフトを入れてウイルス定義ファイルを自動更新するなどのように，パソコンをウイルスから守るための対策を行っていますか？」という文には「悪質なコードを検出できる」というコード（符号）を割り振っている．この例では，文中に「悪質なコード」などの単語は表れていないが，「ウイルス」が「悪質なコード」を指していると読み取れ，その検出技術の導入を促しているため，このコードが適切であると判断した．

5. 結果

Doc2Vec を用いた提案手法の解析結果は，カラーコード表記（緑：低⇔赤：高）とともに表 1 の左に記載している．

表 1 中の「カテゴリ」の列はカテゴリ毎の文書と解析対象の間の提案手法による解析結果を表し，「機能」の列は機能毎の文書と解析対象の間の結果を表す．「機能ごとのカテゴリ平均」は，「カテゴリ」の解析結果を，機能ごとに平均した値である．

「カテゴリ」と「機能」どちらの場合でも， 全体を一文書としてみなして各カテゴリとのコサイン類似度を計算する a の方法の結果と，解析対象の一文ずつ各項目との類似度を計算して最後に合算した b の方法の結果で，ほぼ同様の傾向が確認できた．

「防御」に関する項目において高い数値を示しており，特に，「情報を保護するためのプロセス及び手順」の項目が高くなっている．個別の項目では，「リスクアセスメント」，「異常とイベント」，「低減」の項目が特に高い値をとっている．

6. 評価

ここでは質的コーディングにより得られた結果との比較を行う．コーティング結果は，尾崎の先行研究のものを用いた．

Doc2Vec を用いた提案手法による結果とコーディングの結果のコサイン類似度を計算することにより，人による分析結果とどの程度差異があるか検討する．

質的コーディングの結果と評価手法の結果では単位が異なるため，正規化を行ってから計算を行う．a の方法では，出力結果は（-1~1）の間を取るため正規化を行う必要がなくそのままコサイン類似度の計算に利用することができる．一方，質的コーディングによるスコアと b.の方法によるスコアは，下記の式 1 と式 2 に基づいて正規化を行いコサイン類似度の計算に用いた．

$$N1(X) = \frac{X - x_{min}}{|x_{max} - x_{min}|} \quad \cdots\cdots式\ 1$$

$$N2(X) = \frac{X}{|x_{max}|} \quad \cdots\cdots式\ 2$$

正規化後のコーディング結果と提案手法による結果のコサイン類似度を，表 2 に記載する．「カテゴリ」は，カテゴリの 23 項目同士でコサイン類似度を計算した場合，「機能」は機能の 5 項目同士でコサイン類似度を計算した場合であり，「機能ごとカテゴリ平均」は，機能ごとのカテゴリの値の平均値を用いてコサイン類似度を計算した結果である．

「カテゴリ」の場合，コサイン類似度は 0.63 程度であり，「機能」の場合，コサイン類似度は 0.7 程度の値となった．一方，「機能ごとカテゴリ平均」を用いた場合は，0.8 程度の結果を得ることができた．尾崎の先行研究の tf-idf による結果は，「カテゴリ」の場合，0.78 程度であり，「機能ごとのカテゴリ平均」の場合，0.94 程度であるため，それよりも 0.1 程度低い精度となってしまった．

また，正規化表現による違いについて検討すると，概ね式 1 による正規化よりも，式 2 による正規化のほうが高い値をとっている．

表 2 コーディング結果とのコサイン類似度による比較

	a. 1文書	b. 1 行毎の総和	
		式1で正規化	式2で正規化
カテゴリ	0.626	0.648	0.644
機能	0.791	0.676	0.775
機能ごとカテゴリ平均	0.831	0.797	0.819

7. 議論と制限

7.1 表 1 で数値の高かった項目について

今回，最も高い数値となった項目が解析対象の文書に由来しない要因で，高くなっている可能性について検討する．

各カテゴリ同士のコサイン類似度について，表 3 に記載する．これによると「防御」の項目のうち，「⑩アクセス制御」，「⑪意識向上及びトレーニング」，「⑫データセキュリ

Vol.2019-CSEC-86 No.66
Vol.2019-SPT-34 No.66
2019/7/24

ティ」,「⑬情報を保護するためのプロセス及び手順」の項目は,他の項目とも高い類似性を保っていることが伺える.

また,「防御」以外の項目では「特定」に関する項目が,他の項目と類似性を持っていることが伺える.

他のカテゴリとの高い類似性があるカテゴリは,類似性の少ないカテゴリに比べて,数値が高くなりやすいと考えられるため,今回「防御」の数値が高く出ているのは,この性質の影響を受けた可能性もあると考えられる.

一方で,「①異常とイベント」,「㉒低減」については,他のカテゴリとある程度類似性を持っているものの,高いとは言えない.実際に,「①異常とイベント」,「㉒低減」の文書を確認しても似ているか似ていないかの判別を行うことは難しいが,コサイン類似度が 0.9 を超えるほど一致しているとは考えにくい(同様に「⑦リスクアセスメント」「⑬情報を保護するためのプロセス及び手順」も 0.9 を超えるコサイン類似度を示している).そのため,これらの項目ではほかの要因により高い数値が出ている可能性があるか別途検討が必要である.

7.2 質的コーディングの制限について

本研究では,尾崎の先行研究の質的コーディングの結果を評価データとして利用したが,このデータは,質的コーディングにより作成されており,以下の注意が必要である.

1. テンプレートコーディングでよく行われる複数人でのコーディングの実施と統計的なすり合わせ処理は,実施されていない.
2. 提案手法の試行前に質的コーディングを実施され,コーディング結果についてもレビューを行われた.
3. Cybersecurity Framework1.1 の 106 項目のサブカテゴリをコード群とした.これは適切なコードの数より多いと考えられる.しかし,評価段階では,カテゴリに集約して利用しており,フレームワークコアにおいてサブカテゴリとカテゴリは包括的な関係にあると考えられるので,同一カテゴリ内で発生するレベルでのコードの割り当てミスや解釈違いについては,評価結果への影響はないと考えられる.

8. 結論

本研究では,「情報セキュリティに関連するガイドラインの内容提示の手法の提案とその評価」の文書内容を提示する手法において,Cybersecurity Framework 1.1 の「フレームワークコア」以外のフレームワークを用づいて,Doc2Vec を用いて解析を行い,質的コーディングの情報をもとに評価を行った.

その結果,カテゴリでの評価(項目数 23)の場合,0.6 程度のコサイン類似度を得ることができ,機能での評価(項目数 5)の場合,0.7 程度のコサイン類似度を得ることができた.

これは,先行研究の tf-idf による結果よりも,0.1 ほど低くなっている.コサイン類似度が低い直接の原因としては,「①異常とイベント」,「㉒低減」,「⑦リスクアセスメント」,「⑬情報を保護するためのプロセス及び手順」の項目が突出して高い数値を示しているためと考えられる.しかしながら,これらのカテゴリが,特に他のカテゴリと類似しているわけでもなく,他の要因により,高いコサイン類似度

表 3 各カテゴリ同士のコサイン類似度

		①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰	⑱	⑲	⑳	㉑	㉒	㉓
検知	①異常とイベント	0.98	0.35	0.13	0.49	0.43	0.37	0.39	0.33	0.34	0.44	0.40	0.50	0.46	0.22	0.40	0.15	0.11	0.12	0.31	0.42	0.13	0.21	0.06
	②セキュリティの継続的なモニタリング	0.32	0.98	0.18	0.44	0.45	0.38	0.38	0.36	0.41	0.55	0.49	0.44	0.55	0.17	0.44	0.24	0.18	0.14	0.50	0.47	0.22	0.32	0.09
	③検知プロセス	0.16	0.20	0.99	0.23	0.23	0.26	0.20	0.21	0.28	0.26	0.25	0.27	0.34	0.23	0.22	0.18	0.04	0.09	0.21	0.24	0.08	0.16	0.03
特定	④資産管理	0.42	0.47	0.26	0.98	0.63	0.48	0.49	0.50	0.49	0.69	0.56	0.71	0.74	0.24	0.62	0.29	0.15	0.10	0.52	0.59	0.26	0.41	0.13
	⑤ビジネス環境	0.45	0.43	0.25	0.67	0.98	0.50	0.53	0.50	0.54	0.69	0.64	0.71	0.71	0.32	0.61	0.37	0.10	0.06	0.54	0.63	0.23	0.42	0.15
	⑥ガバナンス	0.33	0.41	0.25	0.48	0.49	0.98	0.51	0.33	0.39	0.51	0.45	0.57	0.60	0.21	0.50	0.22	0.09	0.17	0.45	0.53	0.34	0.34	0.09
	⑦リスクアセスメント	0.34	0.37	0.24	0.51	0.53	0.46	0.98	0.37	0.50	0.58	0.49	0.64	0.65	0.28	0.50	0.34	0.08	0.00	0.45	0.47	0.33	0.37	0.12
	⑧リスク管理戦略	0.30	0.35	0.20	0.46	0.49	0.28	0.37	0.98	0.45	0.50	0.50	0.48	0.57	0.25	0.46	0.25	0.06	0.02	0.42	0.48	0.22	0.30	0.12
	⑨サプライチェーンリスク	0.31	0.42	0.22	0.49	0.53	0.41	0.45	0.41	0.98	0.61	0.52	0.59	0.62	0.17	0.43	0.28	0.09	0.01	0.53	0.58	0.23	0.35	0.05
防御	⑩アクセス制御	0.45	0.56	0.32	0.69	0.69	0.51	0.62	0.51	0.62	0.98	0.71	0.77	0.82	0.37	0.65	0.34	0.12	0.13	0.61	0.69	0.34	0.39	0.18
	⑪意識向上およびトレーニング	0.39	0.46	0.24	0.58	0.61	0.41	0.51	0.51	0.49	0.69	0.98	0.68	0.72	0.31	0.62	0.29	0.16	0.06	0.53	0.60	0.35	0.39	0.15
	⑫データセキュリティ	0.49	0.55	0.34	0.69	0.69	0.57	0.65	0.51	0.59	0.80	0.70	0.98	0.87	0.27	0.64	0.29	0.12	0.11	0.64	0.72	0.34	0.46	0.23
	⑬情報を保護するためのプロセスおよび手順	0.51	0.61	0.29	0.73	0.71	0.57	0.68	0.51	0.60	0.84	0.67	0.86	0.99	0.38	0.69	0.45	0.14	0.10	0.69	0.75	0.39	0.46	0.11
	⑭保守	0.25	0.19	0.24	0.24	0.31	0.24	0.30	0.30	0.19	0.40	0.34	0.32	0.34	0.99	0.25	0.15	0.05	-0.02	0.23	0.31	0.15	0.16	0.13
	⑮保護技術	0.40	0.43	0.21	0.59	0.62	0.52	0.53	0.45	0.43	0.63	0.62	0.66	0.66	0.24	0.98	0.35	0.11	0.12	0.53	0.66	0.29	0.34	0.07
復旧	⑯改善	0.15	0.26	0.22	0.29	0.29	0.22	0.34	0.25	0.28	0.37	0.30	0.30	0.39	0.18	0.25	0.99	0.10	0.06	0.26	0.32	0.17	0.20	-0.03
	⑰コミュニケーション	0.11	0.17	0.02	0.14	0.09	0.08	0.06	0.04	0.09	0.11	0.17	0.08	0.19	0.08	0.10	0.12	1.00	0.05	0.10	0.14	0.14	-0.01	-0.02
	⑱復旧計画	0.12	0.10	0.10	0.10	0.07	0.16	0.01	0.01	0.02	0.10	0.06	0.16	0.08	-0.02	0.12	0.03	0.06	1.00	0.04	0.06	-0.02	0.07	-0.04
対応	⑲分析	0.31	0.47	0.21	0.54	0.53	0.46	0.46	0.42	0.47	0.57	0.51	0.72	0.68	0.26	0.54	0.29	0.09	0.04	0.98	0.60	0.36	0.35	0.12
	⑳コミュニケーション	0.39	0.43	0.25	0.56	0.68	0.57	0.47	0.43	0.54	0.72	0.60	0.72	0.75	0.30	0.65	0.31	0.18	0.07	0.56	0.98	0.33	0.35	0.16
	㉑改善	0.15	0.22	0.10	0.27	0.25	0.33	0.30	0.23	0.22	0.28	0.31	0.33	0.31	0.16	0.26	0.22	0.16	-0.03	0.33	0.30	0.99	0.11	0.13
	㉒低減	0.20	0.30	0.18	0.41	0.33	0.36	0.38	0.28	0.39	0.44	0.33	0.33	0.46	0.13	0.34	0.23	0.00	0.08	0.32	0.33	0.11	0.98	0.03
	㉓対応計画	0.09	0.10	0.02	0.11	0.18	0.10	0.14	0.13	0.07	0.20	0.18	0.21	0.15	0.10	0.16	-0.05	-0.02	-0.04	0.10	0.17	0.11	0.00	1.00

を示しているものと考えられ，別途検討が必要だと考えている．

参考文献

[1] “「情報セキュリティ人材の育成に関する基礎調査」報告書について”. https://www.ipa.go.jp/security/fy23/reports/jinzai/, (参照 2019-06-21).
[2] “情報セキュリティ人材不足数等に関する追加分析について（概要）”. https://www.ipa.go.jp/files/000040646.pdf, (参照 2019-06-21).
[3] “サイバーセキュリティ人材の育成に関する施策間連携ワーキンググループ報告書”. https://www.nisc.go.jp/conference/cs/pdf/jinzai-sesaku2018set.pdf, (参照 2019-06-21).
[4] “情報セキュリティ事故に関わるアンケート調査”. http://lab.iisec.ac.jp/~hiromatsu_lab/files/jiko-questionnaire_result.pdf , (参照 2019-06-21).
[5] “法人組織におけるセキュリティ実態調査 2017 年版”. https://appweb.trendmicro.com/doc_dl/select.asp?type=1&cid=236, (参照 2019-06-21)
[6] 中矢 誠, 富永 浩之. Web ゲームサイトを題材とした攻防型ハッキング競技の環境構築と運用実践 - 試行実践に基づいて改善を行った本番実践の結果と分析. 2018,vol 12, 研究報告コンピュータと教育（CE), p1-8
[7] 阿部 隆幸, 中矢 誠, 太田 翔也, 富永 浩之. 学校機関ごとの個別情報を組み込んだ情報セキュリティの導入教育のためのクイズ形式のアドベンチャーゲームの試作. 2017, 第 79 回全国大会講演論文集 p 737 -73
[8] 楠目 幹,阿部 隆幸, 中矢 誠,富永 浩之. 情報セキュリティの導入教育のための大会イベント BeeCon におけるハッキング競技 CTF の問題構築, 2017 第 79 回全国大会講演論文集 p 739 – 740
[9] 湯川 誠人,井口 信和. 仮想マシンを用いた攻防戦型ネットワークセキュリティ学習支援システムにおけるネットワーク型 IDS を用いた不正侵入シナリオの実装, 2018, インターネットと運用技術シンポジウム論文集, 92 – 99
[10] 福山 和生, 谷口 義明, 井口 信和. 仮想マシンを活用したネットワークセキュリティ学習支援システムの実装と評価, 2016, 情報処理学会論文誌, vol. 57, No. 3, p931-935
[11] 孫 英敬, 山口 由紀, 島田創, 高倉弘喜, 谷口 義明.技術能力に注目した情報セキュリティ教育課程開発のためのカリキュラム分析, 2017, 情報処理学会論文誌, vol. 58, No. 5, p1163-1174
[12] ” NICE Cybersecurity Workforce Framework”. https://www.nist.gov/itl/applied-cybersecurity/nice/resources/nice-cybersecurity-workforce-framework (参照 2019-06-21)
[13] 尾崎 敏司. 情報セキュリティに関連するガイドラインの内容提示の手法の提案とその評価. 2019, CE-148, 研究報告コンピュータと教育（CE) , p1-8
[14] ”CYBERSECURITY FRAMEWORK”. https://www.nist.gov/cyberframework, (参照 2019-06-21)
[15] Distributed Representations of Sentences and Documents Proceedings of The 31st International Conference on Machine Learning (ICML 2014), pp. 1188 – 1196
[16] 瀬尾 美紀子. 自律的・依存的援助要請における学習観とつまずき明確化方略の役割−多母集団同時分析による中学・高校生の発達差の検討−. 2007, 教育心理学研究, 55, p 170-183
[17] 松沼 光奉. 学習内容の体制化と図作成方略が現在完了形の学習に及ぼす効果. 2007, 教育心理研究. 55, p 414-425
[18] “重要インフラのサイバーセキュリティを改善するためのフレームワーク 1.1 版”. https://www.ipa.go.jp/files/000071204.pdf, (参照 2019-06-21)
[19] “中小企業の情報セキュリティ対策ガイドライン”. https://www.ipa.go.jp/security/keihatsu/sme/guideline/index.html (参照 2019-03-12)
[20] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto: Applying Conditional Random Fields to Japanese Morphological Analysis, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237
[21]”gensim” https://radimrehurek.com/gensim/