Vol.2013-DPS-154 No.36
Vol.2013-CSEC-60 No.36
2013/3/15

承認コンテキスト類似性を用いた承認誤りリスク評価手法の提案

本多聡美[†1]　綿口吉郎[†1]　大久保隆夫[†1]　金谷延幸[†1]

IT ガバナンスの達成や情報漏洩対策の実施のための手段として，あらかじめ決められたルートに従い申請の内容の許可を依頼する承認プロセスの実行があるが，同種の申請を承認者が多数処理することにより，拒否すべき申請を誤って許可する，あるいはその逆が発生するリスク（以下，承認誤りリスク）が存在する．本稿では，ワークフローエンジンによる管理がなく，依頼者が作成したファイルを承認者へ送付することで許可を依頼する場合における承認誤りリスクを軽減するための制御を目的としたリスク評価手法の提案を行う．提案手法では，ファイルの複製元および承認者の処理履歴を考慮したファイルの類似性（承認コンテキスト類似性）の観点から承認誤りリスクを評価する．情報セキュリティ管理者は，本手法による評価結果を用いて，承認誤りリスクが大きいファイルの処理する際にのみ承認者に対して注意喚起を行うといった，リスクの大小に応じた制御を行うことが可能となる．本手法について，コンテンツ類似性による評価方法と比較するための実験を行った結果，申請書ファイルと同一の語を多く持つ非申請書ファイルを判断対象から排除でき，提案手法の有効性が確認できた．

Security Risk Assessment for Reviewers based on Context Similarity composed of Workflow States and Form Templates

SATOMI HONDA[†1]　YOSHIRO WATAGUCHI[†1]
TAKAO OKUBO[†1]　NOBUYUKI KANAYA[†1]

In workflow management, clients request their application form to their reviewers along predefined routes in order to prevent information leakage or achieve IT-governance. However, there exists a security risk that reviewers mistake the decision in case where they must confirm a lot of kinds of similar forms. We propose an evaluation method for the risk of reviewers' mistakes using context-based similarity composed of workflow states and form templates. Our method aims at relieving security risks when clients send when clients send files made by themselves to their reviewers without workflow engines. An information security manager enables to control reviewers more efficiently according to the result of our proposed method. For example he can alert reviewers only if they must confirm forms that have high risk of mistakes. As a result of comparison of our method with the method using content similarity, we confirmed that we could exclude non-application forms that have several words same as application forms.

1. はじめに

IT ガバナンスの達成や情報漏洩対策の実施のための手段として，あらかじめ決められたルートに従い申請の内容の許可を依頼する承認プロセスの実行がある．依頼者は申請書となるファイル（以下，申請書ファイル）を承認者に送信し，承認者はファイルの内容や依頼者を検証する等して，申請を許可するか否かを判断する．このような承認プロセスの実行は，一般的なセキュリティマネジメントシステムやメール誤送信対策等[1)][2)][3)]に広く利用されている．

しかし，この承認プロセスがより多くのシーンで実行されることより，承認者は様々な依頼者から送信された同種の申請を多数処理しなければならない場面が発生する．そのような場合，承認者が本来拒否すべき申請を許可する，あるいは許可すべき申請を拒否する事態が発生するリスク（以下，承認誤りリスク）も増加すると考えられる．

情報セキュリティ管理者が取れる対策として，承認者に確認したことを示すチェックボックスの記入を指示したり承認者の判断結果を監査したりするといった手段が多数存在するが，承認者が受信するファイルや申請によってこの承認誤りリスクの大小は異なるため，一律な対策を設定したとしても，対策が不十分な個所が発生する，あるいは必要以上の対策を設定することで承認者の利便性が低下する，といった課題がある．

本稿では，複数の n 人の依頼者が依頼者自身の作成したファイルを 1 人の承認者へ送信することで許可を依頼するような，ワークフローエンジンによる管理が存在しない条件の下で，ファイルの複製元および承認者の処理履歴を考慮したファイルの類似性(以下,承認コンテキスト類似性)の観点から承認誤りリスクを評価する手法を提案する．さらに提案手法によりファイルの類似性のみの場合と比較して承認誤りリスクを評価することができることを示す評価実験を行い，申請書ファイルと同一の語を多く持つ非申請書ファイルを判断対象から排除でき，提案手法の有効性が確認できた．

以下，第 2 章で関連技術について述べた後，第 3 章で本稿の対象とする承認プロセスを定義した上で，第 4 章で提案手法，第 5 章で提案手法を利用した承認誤り抑止システムの構成例を述べる．第 6 章で提案手法の有効性について評価実験を行った結果を述べ，第 7 章で評価実験の考察と

†1 (株)富士通研究所
Fujitsu Laboratories Ltd.

Vol.2013-DPS-154 No.36
Vol.2013-CSEC-60 No.36
2013/3/15

提案手法の課題に関する議論を述べる．第 8 章でまとめとする．

2. 関連技術

本章では，セキュリティリスクアセスメント，業務プロセス分析，複製元追跡に関するファイル管理技術を挙げる．

セキュリティリスクアセスメントに関して，資産，脅威，対策をセキュリティ専門家が定量化することで，脅威や対策コストによって失われる資産が最小になるようにセキュリティ対策を選択する手法[4)][5)]や，金利政策におけるインフレと失業率のトレードオフを示すモデルを，USB メモリの使用による利便性向上とセキュリティ低下のトレードオフに応用できることを示す研究[6)]，管理者へのインタビュー結果に基づくセキュリティ資産の分類を行った調査研究[7)]がある．しかし，これらの手法を実際に適用するにはリスクやセキュリティを定量化する必要があるものの，この処理は属人的であるため，同一の条件下でも評価結果が異なるといった課題が存在する．

データベースのアクセス履歴や更新履歴等から業務プロセスを可視化・分析することで業務全体の効率向上を目的とした技術が存在する[8)][9)][10)]．これらの技術により通常とは異なる業務プロセスを発見することで，セキュリティリスクの大きい業務プロセスやイベントを検知することも可能である．また業務プロセス定義の段階でセキュリティ面での目標，それに向けた対策を設定するといった，セキュリティ要素を組み込んだ業務プロセス構築に関する研究がある[11)]．なお本稿では，業務プロセスの分析のみでは検知できないような，通常の業務プロセスの実行におけるセキュリティリスクの評価を目的とする．

ファイル管理技術に関して，組織内に存在するファイルを効率良く管理するための技術として，ファイルの変更や更新，複製等の処理を監視し追跡することができる技術が多数存在する[12)][13)]．本稿では，このようなファイル管理技術を利用し，システムから自動で取得可能な定量的な情報から，承認プロセスの実行における承認誤りリスクの大小を評価するアプローチを取る．

3. 対象とする承認プロセス

3.1 承認操作手順

本稿では n 人の依頼者が，m 種類ある申請の許可を 1 人の承認者へ依頼するような承認手順を対象とする．このとき，図 1 に示すような 3 つの手順から構成される．手順 1 では，依頼者 $_i$ が，申請 $_j$ の許可を承認者へ依頼するとき，まず申請 $_j$ に対応する申請書となるファイル（以下，申請書ファイル）を作成する．作成手順としては，申請 $_j$ の雛形ファイルの複製を取得し編集する，あるいは他の誰かが過去に作成した申請$_j$に対応する申請書となるファイル（以下，申請書ファイル）の複製を受け取りその内容を編集することで作成する．手順 2 では，依頼者 $_i$ は申請書ファイルを承認者へ送信することで，申請の許可を依頼する．申請書ファイルを受け取った承認者は，申請書ファイルの内容に記入間違えがないか，依頼者が許可するにふさわしいか等を検証し，申請を許可するか拒否するかを決定する．最後に手順 3 として，承認者は申請書ファイルと拒否の決定結果を承認結果データベースへ登録する（図 1）．

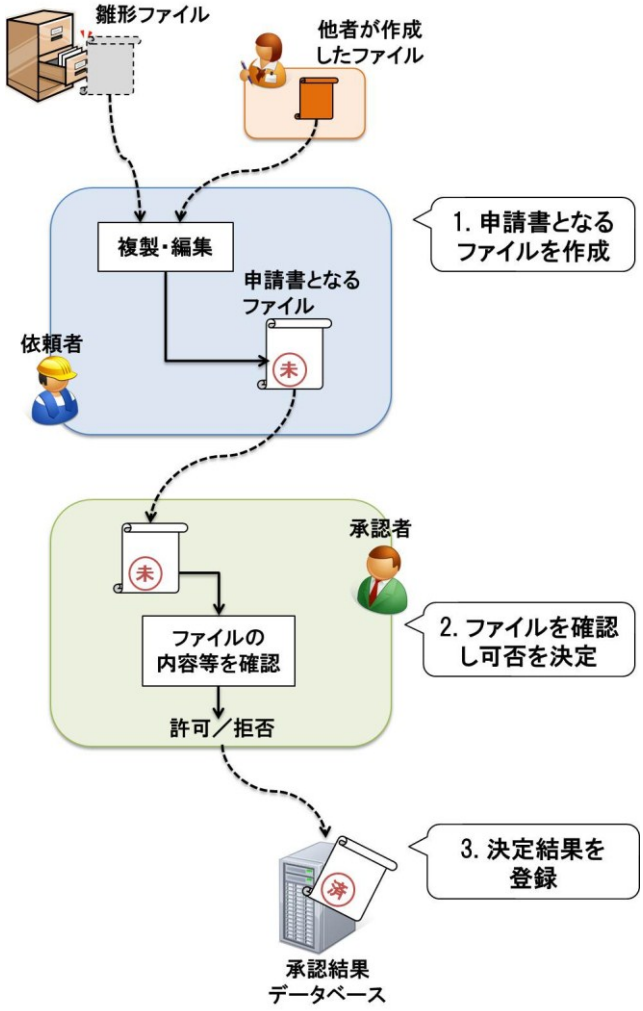

図 1 本稿で対象とする承認プロセス
Figure 1 Workflow process model defined in this paper

3.2 承認プロセスにおけるファイル

本稿ではワークフローエンジンによる管理がない条件下を対象としている．情報セキュリティ管理者は，全ての依頼者が所持する全てのファイルの内容を知ることができ，ファイルの変更や更新，複製等の処理を監視し追跡することもできる．依頼者は申請書ファイルを自身の端末等のローカル環境で作成する．承認者は依頼者から申請書ファイルと非申請書ファイルを受信する．

Vol.2013-DPS-154 No.36
Vol.2013-CSEC-60 No.36
2013/3/15

4. 承認コンテキスト類似性を用いた承認誤りリスク評価手法の提案

本章では，既存のコンテンツ類似性を用いたファイル分類手法における課題を述べた後，コンテンツ類似性を拡張した承認コンテキスト類似性を用いたファイル分類手法を提案し，この提案手法を用いた承認誤りリスク評価手法について述べる．

4.1 コンテンツ類似性を用いた分類手法における課題

第 3 章にて定義した承認プロセスにおいて，プロセスの実行時には承認者は同種の多数の申請書ファイルを多数処理しなければならない場面が発生する．このような場面では承認者は処理に追われる等により，本来許可すべき申請を拒否する，あるいは拒否すべき申請を許可する事態が発生する危険(承認誤りリスク)も大きくなると考えられる．

この「同種の多数の申請書」が存在することを判断する手段として，ファイル中の単語や文章レベルでの類似性判断手法[14)][15)]の利用が考えられる．2 つのファイルが含む共通の単語や表現が多いほどそれら 2 つのファイルの類似度が高いと判断する方法を用いてその類似度によりファイルを分類する．多くの要素を持つ集合に分類されたファイルを承認者が処理する場合には承認誤りリスクが大きくなる，と判断する．この方法を「コンテンツ類似性による判断方法」とする．

しかし，今回想定する条件下においてこのコンテンツ類似性による判断方法を適用する場合には，次の 1～3 に示す課題が存在する（図 2）．

1. **非申請書ファイルを含めた判断**
   申請書として承認者に送信されないファイル（以下，非申請書ファイル）も含めて類似性が判断される．例えば依頼者が独自に作成したファイルに申請書に含まれるような単語が多く含まれていた場合，そのファイルも含めて「類似したファイルが多い」と判断される．
2. **申請の種類を考慮しない判断**
   異なる種類の申請の申請書ファイルが存在した場合に，申請の種類を考慮した判断ができない．たとえ類似する単語が含まれていたとしても，異なる種類の申請を行う申請書ファイルを含むならば，「類似したファイルが多い」ことと「承認誤りリスクが大きい」ことは必ずしも同値ではないと考えられるからである．
3. **送信可能性の低い申請書ファイルを含めた判断**
   処理済みの申請書ファイルも含めて類似性が判断される．例えば承認者に送信され，許可が決定された申請書が内容の変更なしに再び承認者に送信される場合は非常に少ない．そういった既に可否された申請書ファイルも含めて「類似したファイルが多い」と判断される．

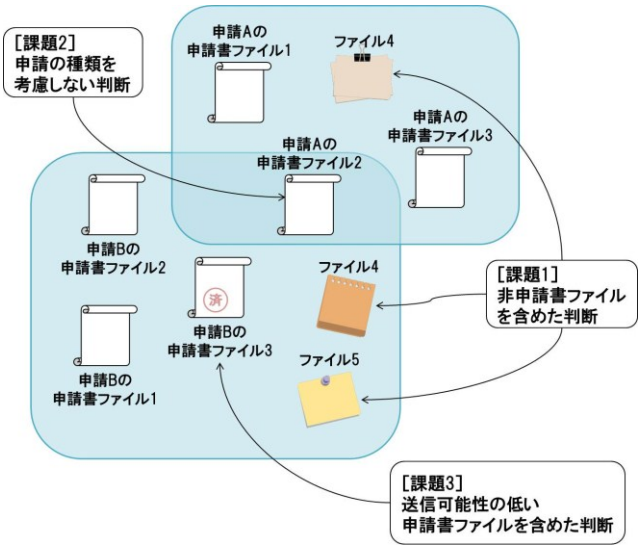

図 2　コンテンツ類似性による判断
Figure 2　Evaluation using contents based similarity

4.2 承認コンテキスト類似性の提案

そこで，承認プロセスにおける状態とファイル複製元から構成されるコンテキスト情報を考慮した承認者にとってのファイル類似性である，承認コンテキスト類似性を提案する．この承認コンテキスト類似性による判断方法では次の 2 点を判断する（図 3）．

- **複製元により申請の種類を特定**…ファイルの複製元を追跡し，そのファイルが雛形ファイルから複製されたかを特定することで，そのファイルがどの種類の申請であるかを判断する．
- **処理の有無を判断**…承認結果データベースにファイルを問い合わせることで，そのファイルが承認者によって可否されたか否かを判断する．ファイルがデータベース内に存在するならば，承認者へ送信される可能性が低いと判断する．

これらの要素を追加することで，コンテンツ類似性と比較して次の 3 点で承認誤りリスクをより正確に判断することができる．

- 非申請書ファイルを判断対象から除外することができる
- 申請の種類ごとに類似性を判断することができる
- 承認者へ送信される可能性の低いファイルを判断対象から除外することができる

Vol.2013-DPS-154 No.36
Vol.2013-CSEC-60 No.36
2013/3/15

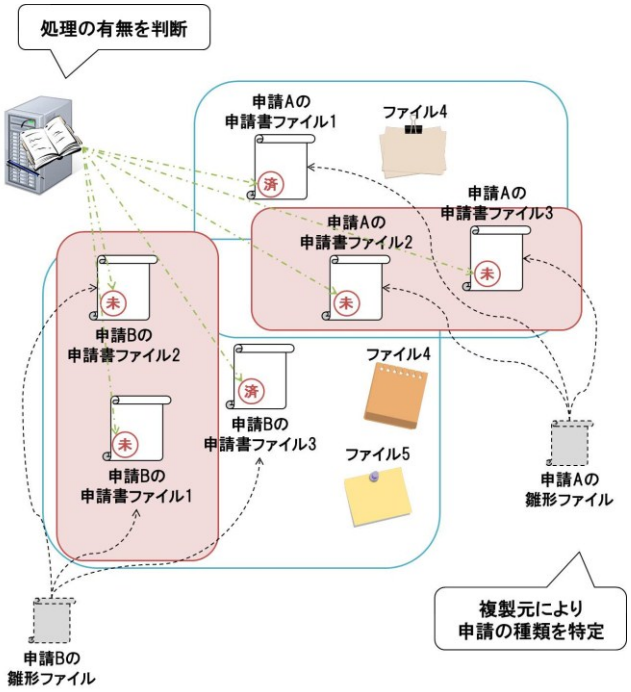

図 3　承認コンテキスト類似性による判断
Figure 3　Evaluation using workflow-context based similarity

4.3 提案手法による承認誤りリスク評価手法

依頼者の所持するファイルを入力とし，承認コンテキスト類似性を利用した承認誤りリスクを次の手順により評価する．

(1) 非申請書ファイルを除外

ファイルの複製元を特定し，雛形ファイルから複製されていないファイルを処理の対象から除外する．

(2) 承認者へ送信可能性の低いファイルを除外

(1)の処理で残ったファイルについて承認結果データベースに可否の結果が登録されているかを問い合わせ，登録されているファイルを対象から除外する．

(3) 申請の種類でファイルを分類

(2)の処理で残ったファイルについて，(1)の追跡結果として得られた雛形ファイルからファイルを申請の種類で分類する．

(4) 申請の種類毎にコンテンツ類似性による分類によりリスクの大小を評価

申請の種類毎にコンテンツ類似性によりファイルを分類する．分類されたファイルが属する集合の要素数が多ければ多いほど，承認誤りリスクが大きいと評価する．

5. 提案手法を用いた承認誤り抑止システム

前章にて提案したファイル分類手法による承認誤りリスクの評価結果を利用した承認誤り抑止システムの構成例を図 4 に示す．

まず，依頼者の端末に依頼者の所持するファイルを監視するシステム（ファイル監視システム）を配置し，ファイル情報を承認誤りリスク評価システムへ送信する．

次に，ファイル監視システムから受信したファイル情報を用いて承認コンテキスト類似性を判断する．制御ルール一覧には，あらかじめ類似性の大小に対応した制御手段（ポップアップを表示させる，チェックボックスを表示させる等）が記録されており，類似性判断結果と制御ルール一覧から，どの申請書ファイルにどの制御手段を適用すればよいかを紐づけた情報（制御対象ファイル情報）を出力する．

制御対象ファイル情報は承認者の端末で動作する承認操作監視システムに送信され，承認者が受信した申請書ファイルに対して，制御対象ファイル情報に沿った制御手段が実行される．

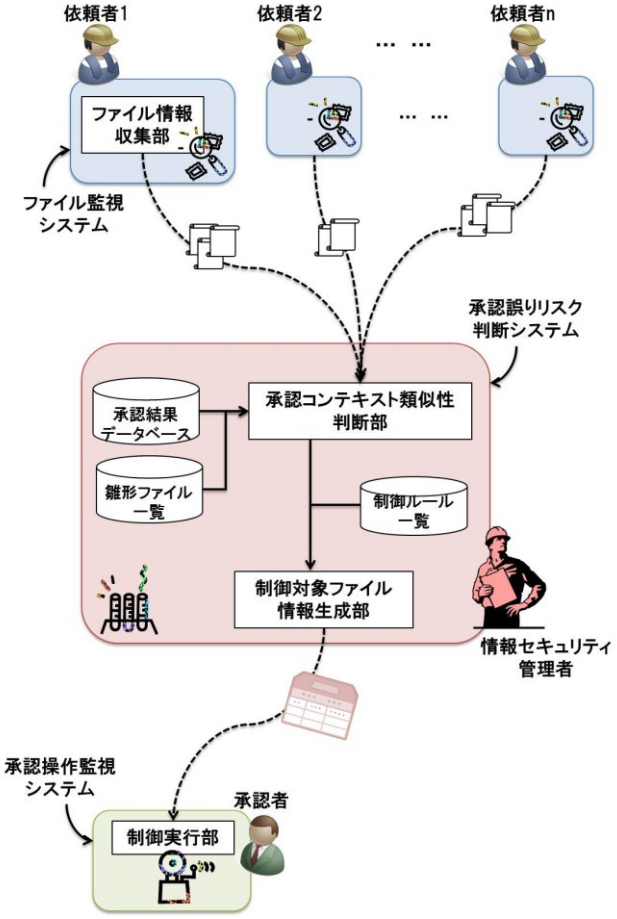

図 4　承認コンテキスト類似性を用いたシステム構成
Figure 4　Architecture using workflow-context similarity

6. 評価実験

コンテンツ類似性のみを利用した場合と比較して，承認コンテキスト類似性を利用する場合により正確に承認誤りリスクを評価することができることを示すため，次に示す要領で評価実験を行った．

6.1 概要

3 種類の申請について，著者の所属する研究グループに所属する 5 人の所持する Microsoft(a) Office Word, Excel[16)]ファイルを対象として，①社外発表申請書，②社外講師派遣申請書，③特許出願申請書についてコンテンツ類似性により分類された集合（コンテンツ集合）と，承認コンテキスト類似性により分類された集合（コンテキスト集合）の 2 種類の集合を作成し，集合に含まれるファイル数の差により評価を行った．

6.2 評価手順

次の手順により①社外発表申請書，②社外講師派遣申請書，③特許出願申請書の 3 種類の申請について，コンテンツ集合，コンテキスト集合を作成した．

(1) コンテンツ集合を作成

各々の申請に対応する申請書ファイルが本文に含むような検索キーワード（表 1）を作成し，Hyper Extraier[17)]を用いてファイル本文を走査した．該当するファイルをコンテンツ集合とした．なお，検索キーワードは関係者の意見を基に作成した．

(2) 非申請書ファイルを除外

雛形ファイルから複製したファイルは，ドキュメントプロパティに特定の情報（メタデータ）が格納されている．それを用いて，(1)でコンテンツ集合に該当するファイルについて，①，②，③の申請書であることを示すメタデータが含まれるファイルを，コンテキスト集合とした．

表 1　コンテンツ集合作成時に用いたキーワード
Table 1 Keywords used for creating contents sets

申請の種類	検索に用いたキーワード
①社外発表申請書	（論文 or 査読 or 予稿集）and 社外発表 and 共同
②社外講師派遣申請書	（大学 or 研究機関） and 講師 and 講義 and 依頼 and 承諾
③特許出願申請書	発明 and 公知 and 作用 and 効果

(a) Microsoft は米国 Microsoft Corporation の米国およびその他の国における登録商標です．

6.3 評価結果

各申請に対する，コンテンツ集合およびコンテキスト集合の作成結果を表 2 と図 5 に示す．ただし，CNT 数はコンテンツ集合に該当したファイル数，CXT 数はコンテキスト集合に該当したファイル数を示す．

表 2　コンテンツ集合とコンテキスト集合に含まれたファイル数
Table 2　Element of contents-similarity sets and workflow-context-similarity sets

申請の種類	CNT 数	CXT 数	$\frac{CXT\,数}{CNT\,数}\times 100$
①社外発表申請書	255	199	78.04
②社外講師派遣申請書	13	4	30.77
③特許出願申請書	82	78	95.12

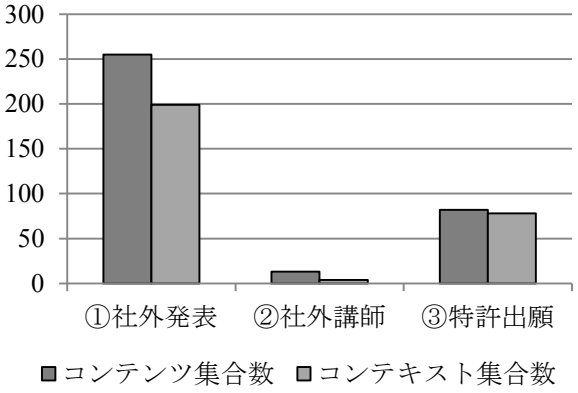

図 5　コンテンツ集合とコンテキスト集合に含まれたファイル数
Figure 5　Elements of contents-similarity and context-similarity sets

コンテンツ集合，コンテキスト集合に含まれたファイルを目視により確認した結果，②社外講師派遣申請書に関しては申請書ファイルを非申請書ファイルと検知したファイルがあった．③特許出願申請書に関しては申請書ファイルの中でも他人が作成した申請書ファイルを複製して所持していた，申請書ファイルを下書きとして複数所持していた，といった事例があった．

7. 議論

7.1 評価実験に関する考察

評価実験の結果，申請書ファイルを下書きとして多数所持している場合には判断結果の誤差が大きくなるものの，コンテンツレベルで類似する非申請書ファイルを，承認コ

ンテキスト類似性による判断手法により排除できたことが確認できた．申請書ファイルを非申請書ファイルと検知したファイルが存在した事例については，雛形ファイルのメタデータを依頼者が編集できないようにするなど，複製元の追跡ができる環境を整えること，他人が作成した申請書ファイルの複製を所持していた事例については，承認結果データベースへ問い合わせる処理を追加することで解決することができると考えられる．

7.2 提案手法における限界

前節までに述べた提案手法における課題として次の 2 つが挙げられる．

まず，例えば下書きを削除せずに残しておくといった，送信する予定のない申請書ファイルが多数存在する場合，実際は承認誤りリスクが低いにもかかわらず，承認コンテキスト類似性による承認誤りリスクは大きいと評価されるため，依頼者の所持する申請書ファイルの中でも，特に送信可能性が高い申請書ファイルを特定するような機能が必要である．

加えて，依頼者が他の依頼者が過去に承認者へ送信した申請書ファイルを受け取り，内容のみを書き換えて承認者へ送信するといった場合には，申請書ファイルが送信される可能性があるにもかかわらず承認コンテキスト類似性の判断対象から除外されるため，例えばファイルの最終編集者や更新日時といった，申請書ファイルの編集履歴も考慮した判断処理も必要である．

8. まとめ

本稿では，複数人の依頼者が依頼者自身の作成したファイルを 1 人の承認者へ送信することで許可を依頼する，ワークフローエンジンによる管理が存在しない条件下において，ファイルの複製元および承認者の処理履歴に基づき承認コンテキスト類似性を判断することによって，承認誤りリスクを評価する手法を提案した．提案手法では，対象とするファイルを限定する機能によりコンテンツ類似性のみによる判断と比べて精度を向上させることが可能である．また本手法について評価実験を行った結果，コンテンツレベルで類似する非申請書ファイルを，提案手法により排除できたことが確認できた．

今後の課題として，今回提案した承認誤りリスク評価結果の検証の実施，さらなる評価精度の向上を目的とした，送信可能性が高い申請書ファイルを特定する機能や申請書ファイルの編集履歴を監視する機能の提案手法への追加が挙げられる．

参考文献

1) http://interstage.fujitsu.com/jp/,"ビジネスアプリケーション基盤 Interstage : 富士通."
2) http://www.hitachi-solutions.co.jp/hibun/sp/product/prod_index03.html,"社外秘持ち出し・セキュリティポリシー｜情報漏洩防止ソリューション 秘文｜日立ソリューションズ."
3) http://www.soliton.co.jp/products/net_security/netattest/filezen/index.html,"セキュア ファイル・データ転送アプライアンス FileZen | TOP."
4) 中村逸一, 兵藤敏之, 曽我正和, 水野忠則, 西垣正勝, “セキュリティ対策選定の実用的な一手法の提案とその評価,” 情報処理学会論文誌 Vol.45 No.8, 2004.
5) 芝口誠仁, 稲場太郎, 中山祐輝, 岡田謙一, “仕事量を考慮したセキュリティ対策選定手法,” 情報処理学会論文誌 Vol.51 No.2, 2010.
6) Adam Beautement, Robert Coles, Jonathan Griffin, Christos Ioannidis, Brian Monahan, David Pym, Angela Sasse, Mike Wonham, “Modelling the Human and Technological Costs and Benefits of USB Memory Stick Security,” Managing Information Risk and the Economics of Security, 2009.
7) Gurpreet Dhillon, Gholamreza Torkzadeh, “Value-focused assessment of information system security in organizations,” Information Systems Journal, 2006.
8) http://bpmcenter.org/,"BPM Center."
9) 日本 IBM, “事例で分かった！ BPM に成功している企業が必ずやっている 3 つのこと,” TechTarget ジャパン ホワイトペーパー, 2011.
10) http://www.questetra.com/ja/,"Cloud Workflow QUESTETRA BPM SUITE."
11) Michael Menzel, Ivonne Thomas, Cristoph Meinel, “Security Requirements Specification in Service-Oriented Business Process Management,” ARES2009, 2009.
12) Walter F. Tichy, “RCS - a system for version control,” Practice and Experience, Vol. 15, No. 7, pp. 637-654, 1985.
13) 福井, 森田, 岡野, 沼尾, 栗原, “ファイルネットワークに基づいた情報の抽出と可視化,” 第 22 回人工知能学会全国大会, 2008.
14) 藤野昭典, 上田修功, 斉藤和巳, "半教師あり学習のための生成・識別ハイブリッド分類器の設計法," 人工知能学会論文誌, Vol.21, No.3, pp.301-309, 2006.
15) 大竹清敬, 増山繁, 山本和英, "名詞の連接情報を用いた関連文書検索手法," 情報処理学会論文誌, Vol.40, No.5, pp.2460-2467, 1999.
16) http://office.microsoft.com/ja-jp/,"Microsoft Office - Office.com."
17) http://fallabs.com/hyperestraier/,"Hyper Estraier: a full-text search system for communities."