

火災調査報告書のテキストマイニングによる出火原因の自動抽出

松場 建都[†] 佐伯 幸郎^{††} 中村 匡秀^{†,††}

[†] 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

^{††} 高知工科大学 〒782-8502 高知県香美市土佐山田町宮ノ口 185

[†] 理化学研究所・革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

E-mail: [†]matsuken@ws.cs.kobe-u.ac.jp, ^{††}saiki.sachio@kochi-tech.ac.jp, ^{†††}masa-n@cmds.kobe-u.ac.jp

あらまし 全国の消防局において、消防常備化に伴い採用された職員の大量退職となる時期を迎えている。益々消防需要が増加するわが国において、若手消防職員への技術・ノウハウ継承は全国の消防行政における問題となっている。神戸市消防局では火災に関するデータを蓄積しており、分析することで防災施策の改善・効率化が期待できる。しかし、火災の詳細な概要が記載されている火災調査報告書は分析が困難であり、十分に活用されていない。そのため、火災調査報告書からの情報抽出が課題として挙げられる。本研究では、火災調査報告書内の火災状況を記載した文章から、出火原因となった発火源、着火物を機械的に抽出することを目的とする。研究のアプローチとして、実際の火災事例で発火源、着火物となった物体を蓄積したインデックスを作成する。その後、作成されたインデックスと火災報告に使用されたデータを活用し、文章から発火源、着火物となる物体の抽出を行う。提案手法を神戸市消防局の所有する火災調査報告書 728 件に対して適用し、発火源を 81.6%、着火物を 74.6% の精度で抽出することができた。

キーワード 技術継承, 火災調査報告書, 出火原因, 自然言語処理, テキストマイニング, 自動抽出

Automatically Extracting the Causes of Fires through Text Mining of Fire Reports

Kento MATSUBA[†], Sachio SAIKI^{††}, and Masahide NAKAMURA^{†,††}

[†] Kobe University Rokkodai-cho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

^{††} Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami City, Kochi 782-8502, JAPAN

[†] Riken AIP, 1-4-1 Nihon-bashi, Chuo-ku, Tokyo 103-0027

E-mail: [†]matsuken@ws.cs.kobe-u.ac.jp, ^{††}saiki.sachio@kochi-tech.ac.jp, ^{†††}masa-n@cmds.kobe-u.ac.jp

Abstract Fire departments across Japan are facing the retirement of a large number of firefighters. So passing on skills and know-how to younger firefighters has become an issue in firefighting administration nationwide. Fire reports, which contain detailed summaries of fires, are difficult to analyze and are not fully utilized. Therefore, extracting information from fire reports is an issue. The purpose of this study is to mechanically extract the ignition sources and ignition materials that caused the fire from the text describing the fire conditions in the fire report. The research approach is to create an index that accumulates objects that have been ignition sources and materials in actual fire cases. Then, utilizing the indexes created and the data used in the fire information, extract the ignition sources and materials from the text. The proposed method was able to extract 81.6% of ignition sources and 74.6% of ignition materials from the fire summary with high accuracy.

Key words know-how transfer, fire report, cause of fire, NLP, text mining, auto extraction

1. はじめに

戦後から現在に至るまで、わが国の消防は市町村における消防の常備化を行ってきた。消防の常備化は昭和 40 年代から急激な進展を続けており、昭和 24 年には 206 本部であった全国の消防本部数は、令和 4 年時点で 723 本部、98.3% の常備化率

となった [1]。このような消防常備化の進展に応じ、多数の消防職員の採用が行われ、令和 5 年現在、常備化の進展に応じて採用された階層の大量退職となる時期を迎えている。高齢化や新型コロナウイルスの影響により、益々消防需要が増加するわが国において、経験豊富な消防職員から若手職員への技術・ノウ

ハウの継承は全国の消防行政における問題となっている [2]。消防行政におけるノウハウ伝承に向けた取り組みの例として、大野らによつて東京消防庁における消防活動経験伝承支援システムが挙げられる [3]。

中でも、ベテランの消防職員が長年の経験に基づいて習得し、データとして残されていない潜在的知識の継承は困難である。潜在的知識の例として、季節性の火災でストーブの可燃物接触による火災が冬場（11月～3月頃）に多く発生する、長期間の保管後の使い初めに、内部に溜まったほこりに着火する、等が挙げられる。このような潜在的知識をデータ化、明文化することにより、若手職員への技術継承や、市民への広報に用いることができる。

潜在的知識の発見、およびデータ化には、火災に関するデータの分析・可視化を通じたアプローチが効果的である。しかし、火災報告データに記載された発火源・着火物は「その他」「ごみ類」などの抽象的な表現が多く、例に示した潜在的知識のような、詳細かつ具体的な結果を得ることは困難である。そこで、火災に関するより具体的な情報を取得するため、火災調査報告書の活用が期待できる。火災調査報告書には火災の出火原因となる発火源、経過、着火物について、火災報告データよりも具体的に記載されており、うまく活用することで火災に関する潜在的知識の発見に役立てることができる。

ところが、火災調査報告書において発火源、着火物が記載されている火災概要文は、自然言語である、書き手によって書式が異なるといった問題があり、機械的な分析は困難である。また、今後増加し続ける報告書すべてに目を通し、人力で発火源、着火物を抽出する作業は消防職員にとって大きな負担となるため、機械的に抽出することが望ましい。

そこで本研究では、火災調査報告書内の火災概要文から、火災原因として重要な発火源、着火物を機械的に抽出する手法を確立することを目的とする。本章のアプローチとして、実際の火災調査報告書から得られる発火源、着火物の情報を蓄積し、それを活用して火災概要文から発火源・着火物の自動抽出を行う。

ケーススタディとして、神戸市消防局の所有する火災調査報告書 728 件に対して提案手法を適用し、発火源、着火物の自動抽出を行った。その結果、発火源を 81.6%、着火物を 74.6% の精度で抽出することができた。

2. 背景知識

2.1 火災と出火原因

火災は消防機関において、「人の意図に反して発生し若しくは拡大し、又は放火により発生して消火の必要がある燃焼現象であって、これを消火するために消火施設又はこれと同程度の効果があるものの利用を必要とするもの、又は人の意図に反して発生若しくは拡大した爆発現象」と定義している [4]。

また消防において、発火源・経過・着火物の 3 つの組み合わせが火災の原因とされており、発火源、経過、着火物の定義は、それぞれ以下のようになっている。

発火源: 出火に直接関係し、またはそれ自体から出火したも

ので、原則として「火」または「高温体」になる。

経過: 火災に至るまでの経過のことで、「現象」・「状態」・「行為」に分類できる。

着火物: 発火源によって最初に着火したもの。

この 3 つの要素に着目して火災に至る経緯を整理、分類し、それに対して適切な予防施策を実施することで、火災の発生を効果的に抑制することができる [5]。

2.2 火災調査報告書

全国の消防局では、すべての火災の原因及び損害並びに関係者の行動等を明らかにし、火災予防施策及び警防対策に必要な基礎資料等消防行政を推進する必要がある。そのため、発生した全火災に対して原因、対応等を調査し、火災調査報告書を作成することが火災調査規程 [6] によって義務づけられている。

火災調査報告書の記載項目の中には、火災の発生日時や、発火源・経過・着火物といった出火原因、その後の延焼物など、火災の概要が簡潔に記載された文章が存在する。本稿では以降これを火災概要文と呼ぶ。火災概要文は、次のような文章である。

この火災の原因は、令和 Y 年 M 月に神戸市 XX 区において、家人が居室で喫煙後、たばこの火を消し忘れたことにより、付近にあったタオルに着火、周囲にあった衣類に延焼したものであると断定する。

火災概要文は、その火災の状況を最も簡潔に、かつ明瞭に表しているものであるが、分析に用いるためには以下の 2 つの問題がある。

- (1) 自然言語であるため分析が困難
- (2) 定まった書式がなく、文章中の発火源・着火物の位置が一定でないため、機械的な抽出が困難

また、火災調査報告書は個人情報を含む文書であるため、分析を外部の機関に委託することは難しい [7]。

火災調査書のテキスト分析の事例として、岡山県岡山市消防局の取り組み [8] や、高田らの防火意匠の現状調査や火災調査書類による延焼分析 [9] などが挙げられるが、いずれの事例でも火災調査報告書からの情報抽出は人の手によって行われており、文章分析においても OCR によるテキスト化にとどまっているのが現状であり、火災概要文からの機械的な情報抽出や得られた情報の分析などは十分に行われていない。

2.3 火災報告

全国の消防において、消防組織法第 22 条の規定に基づき総務省消防庁が各消防本部に求める消防関係報告のうち、火災に関する統計及び情報として「火災報告取扱要領」が定められ、その内容や報告手順等が定められている [10]。要領に基づき、発生時刻や出火箇所、出火原因、損害規模といった火災に関する様々な項目が基礎データとなっている。この報告等を行うための基礎データを基に、その他各消防本部ではデータを作成しており、これを火災報告データと呼ぶこととする。火災データのサンプルを表 1 に示す。火災報告データは、各市の火災に関する統計報告や、市民に向けた広報・注意喚起等のほか、各種分析に活用されている [11]。

表 1: 火災報告データのサンプル

No	出火日時 (年月)	火災種別	火元焼損程度	発火源	経過	着火物	統計上の原因	...
1	2020/01	建物	ぼや	ガスこんろ	不適の用に用いる	繊維製品	ガスコンロ	
2	2020/01	車両	部分焼	たばこ	引火する	不明	不明	

2.4 出火原因分類表

火災の出火原因である発火源、経過、着火物は、火災報告取扱要領にて定められた出火原因分類表に基づいて分類される [10]。分類には大分類、中分類、小分類が存在し、分類ごとの分析、統計報告等に用いられる。

3. 提案手法

本研究では、実際の火災で発火源・着火物として記載されている物体をインデクスとして蓄積し、それをを用いて火災概要文から発火源・着火物を抽出する手法を提案する。実際の火災において、発火源はそれ自身が高温となるものである、着火物は燃えやすい物体であるという特徴がある。このことから、発火源・着火物として抽出すべきフレーズは無限に存在するわけではないため、辞書を用いたキーワード抽出が有効である。そのため、実際の火災概要文に発火源・着火物として記載された名詞を手動で抽出し、蓄積された発火源インデクス及び着火物インデクスを作成する。そして、発火源・着火物インデクスに登録された単語を火災概要文中から探索することで、発火源・着火物の抽出を行う。

火災概要文から発火源・着火物を抽出するまでの具体的なアプローチを以下の A1, A2 に示す。

A1: 発火源・着火物インデクスの作成

A2: 火災概要文から発火源・着火物の機械的抽出

3.1 A1: 発火源・着火物インデクスの作成

A1 では、形態素解析された火災概要文をもとに、発火源インデクス及び着火物インデクスの作成を行う。手順としては、火災概要文から手動で発火源・着火物を抽出する、抽出された発火源・着火物をもとにインデクスを作成する、という流れで各インデクスが作成される。

初めに発火源および着火物の抽出について説明する。発火源、着火物を手動で抽出する際は、誰が抽出作業を行っても同様の抽出結果が得られることが望ましい、そこで、発火源、着火物抽出の条件を以下のように定める。

- 発火源、着火物の抽出数は各火災概要文につき 1 種類とする。
- 抽出者によって異なる物体を発火源、着火物として抽出される事態を回避するため、抽出の際は同火災の火災報告データを参照し、火災報告データに記録された発火源、着火物に適合するものを選択するようにする。火災報告データに発火源が「枯れ草焼き」と記録されている場合、火災概要文中に「火」または「高温体」となる物体が記載されていないケースが多いため、抽出結果を「枯れ草焼き」として処理する。

- 火災概要文の中には、発火源や着火物が記載されていない場合がある。このとき、火災報告データに登録された発火源、着火物は「不明」であることが多いが、例外もある。この場合、火災報告データのデータにかかわらず、発火源、着火物を「記載なし」とする。

発火源、及び着火物の抽出後は、抽出した単語を発火源・着火物インデクスに登録していく。発火源・着火物インデクスの構造を以下に示す。

id: 通し番号。

word: 発火源、及び着火物として抽出された単語。

small_category: 発火源、及び着火物の小分類。

middle_category: 発火源、及び着火物の中分類。

large_category: 発火源、及び着火物の大分類。

frequency: その発火源、及び着火物が追加された回数。

length: 発火源、及び着火物として抽出された単語の文字数。

発火源および着火物の分類に関する項目は、抽出精度向上のためのインデクスのスライスに用いられる。発火源および着火物の追加回数、単語の文字数の項目は、抽出精度向上のためのインデクスのソートに用いられる。

次に、 y 年 m 月に発生した火災 F を例に、抽出した発火源または着火物 w を発火源インデクス、及び着火物インデクスに登録する手順を以下に示す。また、発火源インデクスに単語が登録される流れの一例を図 1 に示す。

手順 1. 火災 F に対応する火災報告データから、発火源小分類または着火物小分類 $s.c$ を取得する。

手順 2-1. 発火源インデクス、または着火物インデクスを参照し、 $word == w, small_category == s.c$ である行 r が既に存在する場合、行 r の $frequency$ の値を 1 増加する。

手順 2-2. 発火源インデクスを参照し、 $word == w, small_category == s.c$ である行 r が存在しない場合、 $frequency = 1$ として発火源インデクスに新たに登録する。

手順 3. 発火源または着火物 w の中分類 $middle_category$ 、および大分類 $large_category$ は出火原因分類表を参照して決定する。

3.2 A2: 火災概要文から発火源・着火物の機械的抽出

A2 では、A1 で作成された発火源インデクス、及び着火物インデクスを活用し、火災概要文から発火源、及び着火物の機械的抽出を行う。辞書に登録された単語を文章中から機械的に抽出する作業はさほど困難でない。本手法においては、火災概要文中にインデクスに登録された単語が複数存在する場合に、どれを適切な発火源および着火物として抽出するかが技術的なチャレンジとなる。本手法では、インデクスの参照順を工夫することで火災概要文から適切な発火源、および着火物の抽出を行う。本手法の主な工夫点は次の 2 点である。

id	word	small category	middle category	large category	frequency	length
1	ライター	ライター	たばことマッチ	火種	20	4
2	たばこ	たばこ	たばことマッチ	火種	29	3
3	ガスコンロ	ガスこんろ	都市ガスを用いる移動可能な道具	ガス油類を燃料とする道具装置	30	5

↓ 発火源を表す単語 : 「ライター」
 発火源小分類 : 「たばことマッチ」

id	word	small category	middle category	large category	frequency	length
1	ライター	ライター	たばことマッチ	火種	21	4
2	たばこ	たばこ	たばことマッチ	火種	29	3
3	ガスコンロ	ガスこんろ	都市ガスを用いる移動可能な道具	ガス油類を燃料とする道具装置	30	5

↓ 発火源を表す単語 : 「石油ストーブ」
 発火源小分類 : 「石油・ガソリンストーブ」

id	word	small category	middle category	large category	frequency	length
1	ライター	ライター	たばことマッチ	火種	20	4
2	たばこ	たばこ	たばことマッチ	火種	29	3
3	ガスコンロ	ガスこんろ	都市ガスを用いる移動可能な道具	ガス油類を燃料とする道具装置	30	5
4	石油ストーブ	石油・ガソリンストーブ	油を燃料とする移動可能な道具	ガス油類を燃料とする道具装置	1	6

図 1: 発火源インデックスの作成例

- category に基づくインデックスのスライス
- frequency, length に基づくインデックスのソート

第一の工夫として、インデックスの項目である小分類、中分類、大分類を活用し、参照するインデックスのスライスを行う。火災事案 1 件に対し、火災概要文のほかに火災報告データが製作されており、発火源および着火物の小分類を知ることができる。そのため、各インデックスに登録されている単語の中から、発火源および着火物の小分類と一致する単語のみを予め抽出することで、火災概要文からの不適切な単語の抽出を防ぎ、抽出精度を高めることが可能となる。

第二の工夫として、インデックスの項目である frequency, length を用いてインデックスのソートを行う。発火源および着火物は火災概要文 1 件につき 1 単語ずつの抽出となるため、インデックスを参照する順番が重要となる。もし発火源や着火物として適切でない単語を概要文から探索・発見した場合、誤った出火原因を抽出してしまうため、発火源、着火物である可能性の高い単語から探索していく必要がある。そこで、インデックス内の単語の参照順を、インデックスの項目である frequency, length の降順とする。frequency の降順に単語を探索することで、他の火災事例で発火源、着火物となることの多い物体を優先的に抽出できる。

発火源抽出の具体的な手順を以下に示す。なお、着火物の抽出手順も発火源と同様である。

- 手順 1. 発火源インデックス項目 length の値で降順にソートする。
- 手順 2. 発火源インデックス項目 frequency の値で降順にソートする。
- 手順 3. 発火源インデックスを発火源小分類の値でスライスし、word が火災概要文中に含まれているか順に探索を行う。word

id	word	small_category~large_category	frequency	length
1	ライター	...	29	4
2	たばこ	...	21	3
3	ガスコンロ	...	30	5
4	石油ストーブ	...	13	6
5	マッチ	...	21	3
6	電気ストーブ	...	6	6

↓ lengthの値で降順ソート

id	word	small_category~large_category	frequency	length
4	石油ストーブ	...	13	6
6	電気ストーブ	...	6	6
3	ガスコンロ	...	30	5
1	ライター	...	29	4
2	たばこ	...	21	3
5	マッチ	...	21	3

↓ frequencyの値で降順ソート

id	word	small_category~large_category	frequency	length
3	ガスコンロ	...	30	5
1	ライター	...	29	4
2	たばこ	...	21	3
5	マッチ	...	21	3
4	石油ストーブ	...	13	6
6	電気ストーブ	...	6	6

図 2: 発火源インデックスのソート例

を火災概要文中に見つけた場合、word を発火源として探索を終了する。

手順 4. 発火源インデックスを発火源中分類の値でスライスし、word が火災概要文中に含まれているか順に探索を行う。word を火災概要文中に見つけた場合、word を発火源として探索を終了する。

手順 5. 発火源インデックスを発火源大分類の値でスライスし、word が火災概要文中に含まれているか順に探索を行う。word を火災概要文中に見つけた場合、word を発火源として探索を終了する。

手順 1, 手順 2 における、発火源インデックスのソートを行う一例を図 2 に示す。また、「YYYY 年 MM 月に神戸市 XX 区において、家人が居室で喫煙後、たばこの火を消し忘れたことにより、付近にあったタオルに着火、周囲にあった衣類に延焼した。」という火災概要の例文から、手順 3 から手順 5 を用いて発火源を抽出する流れを図 3 に示す。

4. ケーススタディ

本手法で火災概要文から正しく発火源・着火物が抽出できるかどうか評価するため、神戸市消防局の所有する火災概要文と火災報告データを活用し、発火源・着火物インデックスの作成、および火災概要文から発火源、着火物の自動抽出を行い、正解率の測定を行う。使用するデータとして、神戸市消防局が 2020 年および 2021 年に出動した火災全 728 件に対し、作成された火災概要文および火災報告データを使用する。728 件の火災概要文を訓練データとテストデータに分割し、訓練データを用いて発火源インデックスおよび着火物インデックスの作成を行う。その後、作成された発火源インデックスおよび着火物インデックスを用いて、テストデータである火災概要文から発火源、着火物の抽出を行う。さらに、抽出された発火源、着火物と、事前に手動で抽出した発火源、着火物の比較を行い、抽出精度の評価

YYYY年MM月に神戸市XX区において、家人が居室で喫煙後、たばこの火を消し忘れたことにより、付近にあったタオルに着火、周囲にあった衣類に延焼した。

発火源小分類：たばこ、発火源中分類：たばことライター

id	word	small_category~large_category	frequency	length
3	ガスコンロ	...	30	5
1	ライター	...	29	4
2	たばこ	...	21	3
5	マッチ	...	21	3
4	石油ストーブ	...	13	6
6	電気ストーブ	...	6	6
:	:	:	:	:
:	:	:	:	:

↓ 発火源小分類「たばこ」でフィルタ

id	word	small_category	middle_category, large_category	frequency	length
2	たばこ	たばこ	...	21	3
9	タバコ	たばこ	...	6	3
14	吸い殻	たばこ	...	4	3
:	:	:	:	:	:
:	:	:	:	:	:

↓ 順にwordを値を文章から探索

id	word	small_category	middle_category, large_category	frequency	length
2	たばこ	たばこ	...	21	3
9	タバコ	たばこ	...	6	3
14	吸い殻	たばこ	...	4	3
:	:	:	:	:	:
:	:	:	:	:	:

↓ 文章内の「たばこ」を発火源として抽出

図 3: 火災概要文から発火源を抽出する例

を行う。各インデックスの作成および発火源、着火物の作成は、Python を用いて行った。

4.1 発火源・着火物インデックスの作成

まず、発火源インデックスおよび着火物インデックスの作成を行う。発火源インデックス作成において、728 件の火災概要文のうち、発火源抽出が不可能である以下の条件を満たす火災概要文を除外した。

- 発火源小分類が「不明」である火災
- 発火源が火災概要文中に記載されていない火災
- 発火源小分類が「枯れ草焼き」である火災

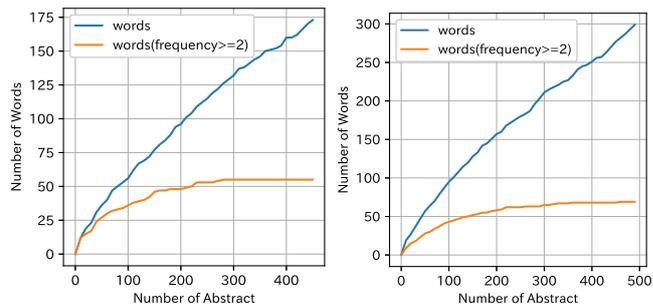
その後、上記の条件を満たさない 570 件を、訓練データ 456 件、テストデータ 114 件に分割し、訓練データ 456 件を用いて、3.1 の手順で発火源インデックスの作成を行った。

また、着火物インデックス作成において、728 件の火災概要文のうち、着火物抽出が不可能である以下の条件を満たす火災概要文を除外した。

- 着火物小分類が「不明」である火災
- 着火物が火災概要文中に記載されていない火災

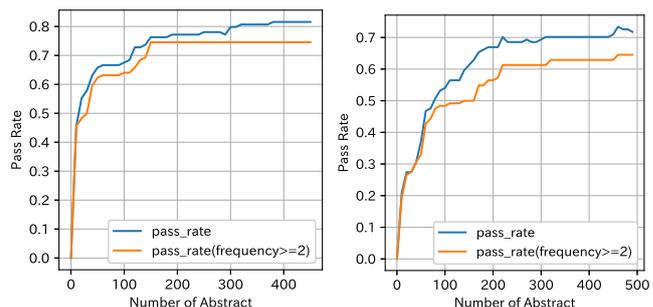
その後、上記の条件を満たさない 620 件を、訓練データ 496 件、テストデータ 124 件に分割し、訓練データ 496 件を用いて、3.1 の手順で着火物インデックスの作成を行った。

各インデックス作成時、火災概要文 10 件ごとにインデックスに登録された単語数、および出現回数 (frequency) が 2 回以上の単語数を記録した。横軸をインデックス作成に使用した火災概要文の件数、縦軸を作成された発火源インデックスの単語数、および出現回数 (frequency) が 2 回以上の単語数の推移としたグラフ



(a) 発火源インデックス (b) 着火物インデックス

図 4: インデックスに登録された単語数



(a) 発火源 (b) 着火物

図 5: 抽出成功率の推移

を図 4(a) に示す。また、横軸をインデックス作成に使用した火災概要文の件数、縦軸を作成された着火物インデックスの単語数、および出現回数 (frequency) が 2 回以上の単語数の推移としたグラフを図 4(b) に示す。

4.2 火災概要文からの発火源・着火物抽出

4.1 で作成した発火源インデックスおよび着火物インデックスを用いて、火災概要文から発火源、着火物の抽出を行い、正解率の評価を行う。発火源抽出は 4.1 で作成した発火源インデックス及びテストデータとした火災概要文 114 件、着火物抽出は 4.1 で作成した着火物インデックス及びテストデータとした火災概要文 124 件を使用する。発火源、および着火物の抽出は、4.1 で記録した火災概要文 10 件ごとのインデックスを用いて行い、抽出された発火源、および着火物と事前に手動で抽出した発火源、着火物を比較し、正解率の推移の算出・可視化を行う。

横軸を発火源インデックス作成に使用した火災概要文の件数、縦軸を発火源抽出の正解率としたグラフを図 5(a) に示す。また、横軸を着火物インデックス作成に使用した火災概要文の件数、縦軸を着火物抽出の正解率としたグラフを図 5(b) に示す。

5. 評価、考察

第一に、発火源インデックスの作成結果について考察する。図 4(a) より、発火源インデックスに登録される単語数は作成時に使用する火災概要文の数に応じて増加を続け、最終的には訓練データの火災概要文 456 件でインデックスに 173 単語が登録された。単語の中には表記ゆれ(「たばこ」「タバコ」など)が含まれるため、実際の発火源はこれよりも小さい数となる。また、

訓練データの火災概要文 456 件で 2 回以上出現した発火源は、インデクス作成に使用する火災概要文が 280 件の時点で 55 個となり、以降増加することはなかった。

第二に、着火物インデクスの作成結果について考察する。図 4(b) より、着火物インデクスに登録される単語数は作成時に使用する火災概要文の数に応じて増加を続け、最終的には訓練データの火災概要文 496 件でインデクスに 299 単語が登録された。発火源インデクスと比較すると、訓練データの件数に対して登録される単語数が多い。これは、発火源に「火、高温体となるもの」という特徴の縛りがあるが、着火物にはそれがないからだと考えられる。また、訓練データの火災概要文 496 件で 2 回以上出現した着火物は、最終的に 69 件となった。

第三に、発火源の抽出結果について考察する。図 5(a) より、発火源インデクスの作成に使用した火災概要文の件数増加に伴い、発火源抽出の正解率は急激に増加し、火災概要文の件数が 150 件の時点で抽出正解率がほぼ一定となった。発火源インデクス内の全単語を用いて抽出を行った場合、最終的な抽出成功率はテストデータ全体の 81.6% であった。また、発火源インデクス内の出現回数が 2 回以上の単語のみを用いて抽出を行った場合、最終的な抽出成功率はテストデータ全体の 74.6% であった。発火源インデクスの frequency の値で抽出に使用する発火源インデクスの単語数を減らした場合、最終的な抽出成功率に約 5% の差が生じた。

第四に、着火物の抽出結果について考察する。図 5(b) より、着火物インデクスの作成に使用した火災概要文の件数増加に伴い、着火物抽出の正解率は急激に増加し、火災概要文の件数が 300 件の時点でほぼ一定となった。着火物インデクス内の全単語を用いて抽出を行った場合、最終的な抽出成功率はテストデータ全体の 71.8% であった。また、着火物インデクス内の出現回数が 2 回以上の単語のみを用いて抽出を行った場合、最終的な抽出成功率はテストデータ全体の 64.5% であった。着火物抽出の成功率は発火源抽出の成功率よりも小さい値であったが、これは発火源インデクスの単語数と同様、着火物には単語としての縛りがなく、抽出が困難であったからと考えられる。

発火源、着火物のどちらにおいても、発火源、着火物インデクスが充実すると未知の火災概要文に対して高い精度で発火源、着火物を抽出することが可能であった。また、発火源、着火物どちらの抽出の場合でも、インデクス作成に使用する火災概要文の量が一定以上になると、抽出成功率が大きな増加をすることがなくなる。このことから、本手法はインデクス作成に用いるデータ数が膨大でなくとも十分な結果が得られることがわかった。

一方で、発火源抽出においてはテストデータ全体の 28.2%、着火物抽出においてはテストデータ全体の 35.5% が抽出失敗という結果であった。抽出失敗には、(1) 適切な発火源、着火物とは異なる単語をインデクスから選択する、(2) インデクス内に適切な発火源、着火物となる単語が登録されていない、の 2 通りの場合がある。(1) の場合、frequency の値に応じて参照するインデクスに適切にフィルターをかけることで、抽出精度低下

を抑制することができる。図 5(a)、図 5(b) のどちらのグラフでも、インデクスを $frequency \geq 2$ の条件でフィルタをかけた場合、インデクス作成に使用する火災概要文数増加に伴う抽出成功率の減少は見られなかった。(2) の場合、インデクスの単語数を増加させることで、インデクスに単語が未登録である場合を減少させることができる。しかし、抽出成功率の大幅に増加させることは本手法では難しく、インデクスへの単語追加にかかるコストも無視できない。そのため今後の展望として、未知語に対する本手法とは異なるアプローチが今後の検討課題である。

6. ま と め

本研究では、神戸市消防局と共同研究を行い、神戸市消防局が持つ火災調査報告書および火災報告データを用いて、火災調査報告書内の火災概要文から発火源、着火物に機械的抽出を行う手法を提案した。

神戸市消防局の所有する火災調査報告書 728 件に対して提案手法を適用し、火災概要文中の発火源を 81.6%、着火物を 74.6% の精度で抽出することができた。

今後の展望として、より多くの火災調査報告書を活用し、未知語への対応を含めた抽出精度の改善や、抽出した発火源、着火物を活用した消防業務改善のための施策構築が挙げられる。

謝辞 本研究の一部は JSPS 科研費 JP19H01138, JP20H05706, JP20H04014, JP20K11059, JP22H03699, JP19K02973, 特別研究員奨励費 22J13217, および、立石科学技術振興財団の研究助成を受けて行われている。本研究は神戸市消防局との共同研究の一環で行われている。

文 献

- [1] 総務省消防庁 (編), “消防白書 令和 3 年版,” 第 2 章, 日経印刷, 2022.
- [2] 稲山博司, “人口減少社会における今後の消防体制の課題等について,” 自治論文集: 地方自治法施行七十周年記念, pp.1027-1042, 2018.
- [3] 光太郎大野, 祐樹小川, 博彦諏訪, 太田敏澄, “東京消防庁における消防活動経験の伝承を支援する sns の提案,” 情報処理学会論文誌, vol.54, no.1, pp.284-294, jan 2013.
- [4] 生田英輔, 赤坂美咲, “大阪市消防局・大阪市立大学共同研究「大阪市における火災分析」,” https://www.city.osaka.lg.jp/contents/wdu010/digitalbook/syobo/kasaibunseki_r2,20213. (Accessed on 12/16/2022).
- [5] 田村裕之, “電気火災の動向と火災事例,” 安全工学, vol.48, no.6, pp.413-418, 2009.
- [6] “火災調査規定,” https://www1.g-reiki.net/city.kobe/reiki_honbun/k302RG00001396.html. (Accessed on 02/24/2023).
- [7] 村上英明他, “消防行政における情報公開および個人情報保護: 情報開示請求に係る事案の考察,” 福岡大学法学論叢, vol.62, no.1, pp.129-251, 2017.
- [8] 消防庁, “第 4 回 予防業務優良事例表彰,” <https://api.ce-cotoha.com/contents/index.html>, July 2020.
- [9] 高田駿平, 平尾和洋, 山本直彦, “奈良県明日香村飛鳥・奥山大字における防火意匠の現状調査と火災調査書類による延焼分析,” 歴史都市防災論文集, vol.9, pp.41-48, 2015.
- [10] 防災行政研究会, 火災報告取扱要領ハンドブック, 東京法令出版, 2022.
- [11] 佐藤博臣, 村井裕樹, 志田弘二, 栗岡均, “火災統計データに基づく火災拡大危険評価法,” 日本建築学会環境系論文集, vol.71, no.605, pp.1-8, 2006.