

米国原子力発電所の異常事象報告書（LER）に関する テキストマイニングによるデータ分析の試行

A trial analysis of license event reports (LERs) of U.S. nuclear power plants by text mining

新崎 雅志 (Masashi Shinzaki)*¹

山口 浩司 (Hiroshi Yamaguchi)*²

要約 海外原子力発電所で発生した事故・故障には、わが国の原子力発電所の安全性や安定運転の向上に参考となる事象がある。原子力安全システム研究所では、米国を中心に欧州、アジアを含む海外の事故・故障情報を入手して分析評価し、類似の不具合事象が国内の加圧水型軽水炉（PWR）発電所で発生することを防止するための改善提言を行ってきた。

本稿では、米国原子力発電所の異常事象報告（LER：License Event Report）を用い、不具合事象関連の用語などに注目して、テキストマイニングによるデータ分析を試行し、その結果をまとめた。

キーワード 米国原子力発電所, 異常事象報告, テキストマイニング

Abstract Accidents and failures that have occurred at overseas nuclear power plants can serve as a reference for improving the safety and stable operation of nuclear power plants in Japan. The Institute of Nuclear Safety System, Inc. (INSS) has obtained and analyzed information on accidents and failures in the United States and other countries, including Europe and Asia, and has made recommendations for improvements to prevent similar failure events from occurring at pressurized water reactor (PWR) power plants in Japan.

In this paper, we made a trial analysis by text mining of license event reports (LERs) of U.S. nuclear power plants in which we focused on terms related to failure events and other information, and we summarized the results herein.

Keywords U.S. nuclear power plants, license event reports, text mining

1. はじめに

原子力安全システム研究所では、米国を中心に欧州、アジアを含む海外の事故・故障情報を入手して分析評価し、類似の不具合事象が国内のPWR発電所で発生することを防止するための改善提言を行っている。

これら個別の事故・故障情報の分析に加えて、テキストマイニングによるデータ分析すなわち計量テキスト分析を活用することにより、更に有効な教訓・提言等を引き出せる可能性がある。

本稿では、海外の事故・故障情報の分析データを基に、不具合事象関連の用語などに注目して、テキストマイニングによるデータ分析を試行し、その結果をまとめた。

本分析に用いる情報ソースは、①公開情報であること、②国内の商用原子力発電所と型式（PWR およびBWR）が同じであること、③記載内容が充実していることから、米国原子力発電所の異常事象報告（LER）

を採用した。なお、LERは、米国連邦規則10CFR50.73 “License event report system” に定める報告基準に基づき事業者から米国原子規制委員会（NRC）に提出される異常事象報告である。

また、テキストマイニングによるデータ分析用のソフトウェアには、KH Coder（樋口，2019）を用いた。

2. 原子力関連の専門用語の辞書作成

KH Coder では、テキスト型データの中から一般的な単語は自動的に取り出せるが、原子力関連の専門用語は自動的に取り出せない。そのため、まず始めに、機器名や不具合事象名などの辞書を作成した。辞書の作成に当たって、KH Coder を用いて2010年から2018年のLER分析データ（2,779件）より自動的に取り出した単語約24,000語から、出現回数が少ない単語を除き、「原子+炉→原子炉」、「運転+不能→運転不能」などを複合語として指定して、2,400語を抽出・整理

*1 元 (株)原子力安全システム研究所 技術システム研究所 現(公財)原子力環境整備促進・資金管理センター

*2 (株)原子力安全システム研究所 技術システム研究所

した。

なお、原子力関連の専門用語の抽出・整理にはかなりの作業量を要した。

3. どのような不具合事象が多いか

2010年から2018年のLER分析データのうち事象概要、原因および対策を対象に、テキストマイニングによるデータ分析を行い、不具合事象関連の用語を抽出した。表1に、不具合事象関連の抽出語を降順に示す。

「トリップ」、「漏えい」、「故障」、「運転不能」、「喪失」、「逸脱」、「損傷」等の出現回数が多かった。

なお、「漏えい」の出現回数は、「漏洩」、「漏れ」または「リーク」の各出現回数も含め合算している。

4. どのような不具合事象が増えたか

前項で抽出した不具合事象関連の用語を、年別（2010年、2011年 … 2018年）に再抽出した。表2に、年別の不具合事象関連の抽出語を降順に示す。

表1 不具合事象関連の抽出語
[2010~2018年]

出現順位	不具合事象	出現回数
1	トリップ	1,697
2	漏えい	1,036
3	故障	912
4	運転不能	555
5	喪失	516
6	逸脱	481
7	損傷	446
8	振動	303
9	劣化	293
10	失敗	248
11	欠陥	233
12	超過	207
13	不足	200
14	作動不能	189
15	地絡	183

表2 年別の不具合事象関連の抽出語

[2010年]

出現頻度	不具合事象	出現回数
1	トリップ	231
2	漏えい	155
3	故障	139
4	運転不能	70
5	喪失	60
6	損傷	54
7	不良	48
8	不足	43
9	スクラム	39
9	逸脱	39
11	腐食	36
11	劣化	36
13	事故	33
13	超過	33
15	失敗	32

[2011年]

出現頻度	不具合事象	出現回数
1	トリップ	243
2	漏えい	158
3	故障	129
4	運転不能	91
5	損傷	90
6	喪失	72
7	逸脱	67
8	不良	49
9	不備	36
10	振動	34
10	劣化	34
12	接触	33
13	不足	32
14	失敗	31
15	腐食	30

....

[2018年]

出現頻度	不具合事象	出現回数
1	トリップ	142
2	故障	120
3	漏えい	96
4	運転不能	47
5	喪失	45
6	逸脱	36
6	損傷	36
8	不良	27
9	不足	26
10	作動不能	25
11	振動	24
11	劣化	24
13	スクラム	23
14	欠陥	21
14	不備	21

全期間[2010～2018年]と同様に、各年とも出現回数の上位をほぼ「トリップ」、「漏えい」、「故障」、「運転不能」、「喪失」、「逸脱」、「損傷」で占めている。2010年から2018年の9年間の各年において、出現回数の上位については、出現回数は年毎に異なるが、出現順位に大きな変化はなく、特定の不具合事象関連の抽出語に増加傾向や減少傾向は見られなかった。

5. 共起ネットワークの作成

更に、KH Coder を用いて、出現回数上位の「漏えい」と関連する抽出語が共通に出現する関係（共起関係）を円と線で示す共起ネットワークを作成した。結果を図1に示す。

共起ネットワークの作成に先立ち、「漏えい」、「漏洩」、

「漏れ」、「リーク」のいずれかが含まれている不具合事象を集計できるようにコーディングルール・ファイルを作成した。このような用語のばらつきは、発電所で慣例的に使用されていたり、分析担当者による翻訳の仕方の違い等によるもので、分析に当って、ここまで細かく用語の翻訳の仕方をルール化していないことから、テキストマイニングによるデータ分析では、このような手間をかける必要が生じる。

このコーディングルール・ファイルを読み込んで検索した結果、「漏えい」、「漏洩」、「漏れ」または「リーク」を含む不具合事象において、図中に赤太一点鎖線で囲ったように「点検で〇〇配管溶接部からの漏れを確認」、「シール部からのリーク発生」、「漏えい試験実施時の△△弁からの漏えい」などの出現回数が高いことが分かった。

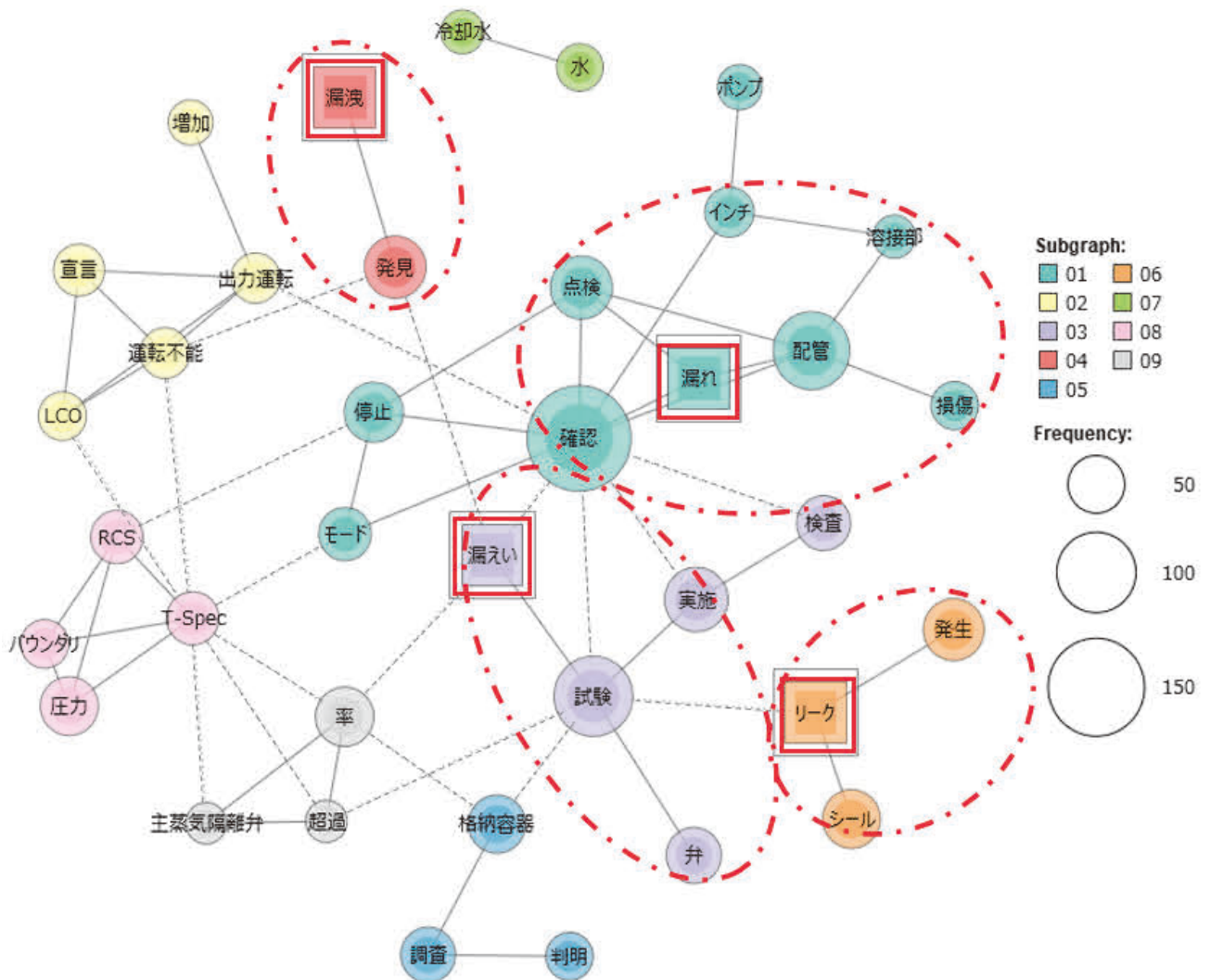


図1 「漏えい」と関連する抽出語との関係を示す共起ネットワーク

6. 機械学習による原因分類

次に、KH Coder を用いた機械学習によって、不具合事象の原因分類を検討させることで、分析者の行う分析評価の作業量の低減につながる方策が得られるのではないかと考えた。

2010年から2018年のLER分析データ(事象の件名、概要、原因および対策に、原因分類、国内PWRプラントへの対策の要否などの分析評価を加えたもの)のうち原因と原因分類を対象に、KH Coder を用いて原因分類のルールを学習させ、学習対象外の原因に対して原因分類を行い、原因分類の一致率を測定した。測定結果を表3に示す。

表3 原因分類の一致率の測定結果

原因分類	原因分類の一致率	異なる分類先
保守不良	60%	設計不良、運転不良、不明
設計不良	63%	保守不良
施工不良	42%	保守不良、製造不良
製造不良	64%	保守不良
管理不良	100%	—
外部事象	100%	—
運転不良	66%	保守不良
不明	43%	保守不良

管理不良および外部事象の原因分類の一致率は100%となったが、保守不良、設計不良、施工不良、製造不良、運転不良および不明の原因分類の一致率は40~60%程度となった。

7. まとめ

海外の事故・故障情報の分析データを基に、不具合事象関連の用語などに注目して、2010年から2018年のLER分析データを対象に、KH Coder を用いたテキストマイニングによるデータ分析を試行した。

その結果、全期間[2010~2018年]と同様に、各年も出現回数の上位をほぼ「トリップ」、「漏えい」、「故障」、「運転不能」、「喪失」、「逸脱」、「損傷」で占めていた。また、出現回数の上位については、出現回数は年毎に異なるが、出現順位に大きな変化はなく、特定の不具合事象関連の抽出語に増加傾向や減少傾向は見られなかった。

また、出現回数上位の「漏えい」について共起ネットワークを作成したところ、「点検で〇〇配管溶接部からの漏れを確認」、「シール部からのリーク発生」、「漏えい試験実施時の△△弁からの漏えい」などの出現回数が高いことが分かった。

更に、機械学習によって原因に対する原因分類を行い、原因分類の一致率を測定したところ、管理不良および外部事象の原因分類の一致率は100%となったが、保守不良、設計不良、施工不良、製造不良、運転不良および不明の原因分類の一致率は40~60%程度となった。

現状では、保守不良、設計不良、施工不良、製造不良、運転不良および不明の原因分類の一致率はそれほど高いとは言えない。

テキストマイニングによるデータ分析を活用することにより、有効な教訓・提言等を引き出したり、分析者の行う分析評価の作業量の低減につながる方策を得るには、漏えい、漏れ等の翻訳のゆらぎを排除するために原文(英単語)を使ってデータ分析するなど、更なる検討が必要と思われる。

また、分析に入る前の原子力関連の専門用語の辞書の作成にかなりの作業量を要しており、分析内容に応じた辞書を作成するなど、専門用語の辞書の作成に当たって作業量がかからないようにするなどの工夫が必要である。なお、作成された辞書は今後の事象分析に大きく役立つと考えられる。

引用文献

- 永山 統啓 (2020). 米国原子力発電所の異常事象報告書 (LER) に関する傾向分析 INSS JOURNAL, Vol. 27, 211.
- 樋口 耕一 (2019). KH Coder 3 リファレンス・マニュアル 2019年9月17日版.