

IEEE論文に基づくIoT研究動向の計量書誌学的調査

藤井 章博

概要

近年、IoT (Internet of Things) に関連したニュースや記事が多く取り上げられており、特に情報工学分野の研究課題として関心が高まっている。本稿は、IoT の概念整理と論文分析に基づく各種関連研究の方向性を整理し、研究計画の立案に資することを目的とする。

分析方法としては、電気・電子分野における世界最大の学会である IEEE の学術論文を利用し、IoT の学術論文年次発表数の推移を調査する。その上で、IoT と結びつきが強いキーワードを抽出する。これらのキーワードは、IoT の研究における応用領域として研究者の関心が高いものと考えられる。基本的な計量書誌学的手法として TF-IDF を利用した。これを基に、論文から得られた文書データベースをクラスタリングすることによって、IoT の応用対象として、私たちの生活にも関係が深く、大きな影響を与えるセキュリティ対策、建築分野などの領域が抽出された。

キーワード : IoT (Internet of Things), survey, IEEE, 計量書誌学, 統計処理, TF-IDF, クラスタリング

1 はじめに

近年、IoT (Internet of Things) に関連したニュースや記事が多く取り上げられている。本稿では論文の内容に対する自然言語処理的な手法を用いて IoT の研究動向を定性的に計量し、研究の動向を把握することを試みる。これによって、科学技術予測調査等の活動における一つの指標を提供したい。

一般に、計量書誌学的研究では、特許や論文など一定の形式に基づく記述がなされた文章をその計測単位として用いる^{2,3)}。本稿では、世界最大の学会である IEEE (Institute of Electrical and Electronic Engineers) の学術論文データベース¹⁾ を利用し、自然言語処理の手法を用い、関連する要素技術として研究者の間で利用頻度が増加しているキーワード「IoT」と関係性の深い他分野のキーワード等を抽出することを試みる。以下では、まず IoT の概念整理を行った上で論文分析の方法の概要とその結果を述べる。

2 IoT とは何か

IoT とは、コンピュータなどの情報・通信機器だけでなく、世の中に存在する様々なモノに通信機能を持たせ、インターネットに接続したり相互に通信することにより、自動認識や自動制御、遠隔計測などを行うことである。既存の調査に基づく、2020 年までに IoT でインターネットに接続するセンサーやデバイスなどの機器の数が 260 億になる見込みで、IoT 関連の製品及びサービスの売上は 3,000 億ドル以上になるとの予想もある。このことから、私たちの生活に大きな影響を及ぼすことがうかがえる。

IoT の定義は必ずしも明確ではないが、従来からこれに近い概念を表現する用語として、「ユビキタス」、「サイバーフィジカル」、「M2M」等がある。

ユビキタスとは、「いつでも、どこでも、だれでも」が恩恵を受けることができるインターフェース、環境、技術のことである。このとき、ユーザー

である人間からみてコンピュータシステムが遍在していることに視点がある。次に、サイバーフィジカル (Cyber Physical) とは、実世界 (Physical System) に浸透した組み込み型のコンピュータシステムなどが構成するセンサーネットワークなどの情報を、サイバー空間 (Cyber System) の情報処理能力と結びつけることを指すようである。この場合は、実世界における情報処理に視点がある。「M2M (Machine to Machine)」は、人を介せずに物と物が通信を行うことを指している。M2M は機械間のみで自動で通信が行われ、そこに人間は一切介入しない。自販機や家電など、機械間のみで自動的に通信が行われるものが M2M である。

IoT とこれらの概念との差異は、IoT の方がより幅広く、通信接続し相互作用する対象として「モノ」を含む点である。このことから、より幅広い活用法によって現実社会に浸透する技術と考えられている。実際に、ユビキタスでは家電やスマートフォンなど人が利用する概念であるのに対し、IoT では衣服や自動車など、必ずしも人を介さずに端末同士、モノ同士が自律連携することが想定されている。さらに、ある定義によると「Things は物理的なものにとらわれず、バーチャルなもの (コンテンツなど) を含む」点が違いであると述べられている¹⁾。

3 IEEE における IoT の 学術論文年次発表数の推移

世界最大の学会である IEEE の論文データベースにおいて、「IoT」を抄録のキーワードに含む論文の

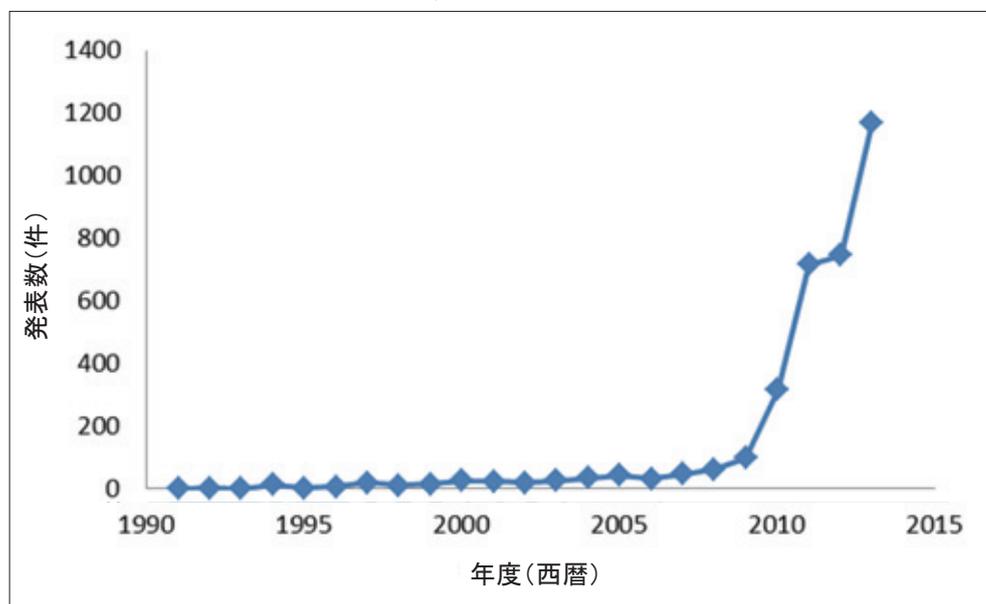
年次発表数の推移を図表 1 に示す。2009 年より急激に学術論文の発表数が増えており、調査対象とした 2013 年末まで 2013 年現在も衰えることなく研究論文数が増加していることが分かる。研究論文の増加に伴い、研究の動向を俯瞰するサーベイ論文も増大する。IEEE 文献データベースでもその傾向は顕著である。そこで、IoT 分野のサーベイ論文の発表状況を調べるため、「IoT」と「survey」の両方のキーワードを持つ論文数を調べたが、同じく 2009 年度より急激に研究が進んでいる。以下では、IoT に関する研究動向を探るために、このサーベイ論文を主たる対象として、計算機を利用した自然言語処理手法による分析を行う。

4 論文中の単語の頻度分析

4-1 頻度分析に使用したサーベイ論文

IoT 分野のサーベイ論文を対象として、自然言語処理によって研究分野の計量分析を行う。はじめに、IEEE 論文データベースを基に、IoT 関連論文の採録 (Abstract) のデータベースを作成した。対象としたのは、IEEE 論文のデータベースから、IoT 分野を俯瞰的に捉えるサーベイ論文の集合を取り出して行う。学会論文誌と国際会議予稿集に掲載された「IoT」に関する論文のうち表題中に「survey」を含み、分析対象となる論文数は 69 本であった。

図表 1 IoT に関する学術論文年次発表数の推移



4-2 高頻度語彙の収集方法とその結果

近年オープンソースプログラムを活用して自然言語処理やテキストマイニングと呼ばれる情報処理が比較的容易に実現できるようになっている³⁾。利用できるツールは様々あるが、本稿では、プログラミング言語 R と併用することができる MTMineR を選んだ⁴⁾。

69 本の論文から、高頻度語彙の収集を行った。処理の概要を以下に列挙する。

- ①データクレンジング：PDF 文書から TXT データ作成、表記の揺らぎ修正等
- ② Abstract のテキストファイルの文書データベースの作成
- ③単語の頻度分析：テキストマイニングツール (MTMineR) を利用した。

69 本の論文に対してデータクレンジングを行い、MTMineR を利用し頻度分析を行った結果を以下に示す。図表 2 には、単語単体の頻度分析の結果の一部を示す。単語の並びの頻度分析の結果の一部を図表 3 に示す。それぞれ、786 個の単語、363 個の単語のリストに対し頻度分析を行った。

図表 2、3 の結果より、internet や networks など、IoT の動向を示すことのできるネットワーク関連のキーワードが多く収集された。特に図表 3 ではセンサーネットワークや RFID など、無線通信関連の技術が多く表れていることが分かる。IoT の動向としては、当然予想できることながら、情報ネットワーク方面と無線通信技術との関わりが強く、そのような文脈の研究が多く実施されていることが分かる。一方、セキュリティやスマートシティ、データマイニングなど、アプリケーションよりの単語も散見される。このことから、IoT は様々な応用分野との関わりを持ち、我々の生活に影響を与えると考

図表 2 単語単体の頻度分析の結果の一部

語彙	頻度
internet	42
sensor	25
networks	20
wireless	20
management	12
social	10
protocol	9
data	8
security	8
energy	7

えられる。

5 クラスタリングに基づく IoT の応用領域の把握

5-1 クラスタリング

クラスタリングとは、データの集合からそれぞれ共通の特徴を持つ複数のグループを抽出・分類する分析手法のことをいう³⁾。分析の対象は、上述した 69 本の IEEE 学術論文を対象とする。クラスタリングを行った上で、それぞれのグループについて基本的な自然言語処理を施すことにより、各グループの特徴を表す単語を抽出し、客観的なデータに基づく動向の把握を行う。ここで利用するのは、TF (Term Frequency) 値が高い単語である。この値の高い単語は、その分野の特徴を相対的に表現する単語と考えられる。一方、文書集合中の特定の文書の特徴は、IDF (Inverse Document Frequency) 値によって表現される。これは、その文書集合の中で特定の文書単体の特徴をよりよく表す単語である。これらの二つの評価値を掛け合わせた TF-IDF 値の高い単語群によりベクトル評価基準を作る。それぞれの文書をこのベクトル量で評価することで、文書のグループへの分類やそのグループの特徴を表すキーワードを把握することができる。以下では、この方法に基づいてサーベイ論文をグループに分類しそれぞれの特徴を把握する。

以下に手順の概略を示す。

- ①文書特徴ベクトル：出現頻度 (TF: Term Frequency) の高い語を用いて特徴ベクトルを構成する³⁾。
- ②ワード法より 6 つのグループに分類した。

図表 3 単語の並びの頻度分析の結果の一部

語彙	頻度
Internet-of-things	33
Wireless-sensor-networks	9
Radio-frequency-identification	7
6lowpan	4
Constrained-application-protocol	4
Smart-cities	4
Social-network	4
Data-mining	3
Human-computer-interaction	3
Middleware	3

- ③各グループの単語に関する頻度分析を行う。
- ④文書の特徴をよく表現する単語 (IDF: Inverse Document Frequency の高い単語) を抽出する。6グループの特徴を表す語を抽出した (図表4)。図表4は、6つのグループそれぞれにおいて抽出されたキーワードを示す。それぞれのグループの特徴を表現する単語が抽出されていることが分かる。IoT 研究は、医療分野、産業分野、サービス業界など、多くの業界と関わりがあることがうかがえる。このことから、IoT は今後もより多くの業界と関係を持ち、私たちの生活に大きく関わるものがうかがえる。

5-2 各クラスターの内容

6つのクラスターの論文の内容に従って、それぞれの特徴を見ていく。論文の内容に加えて Web 上で公開されている情報も関連付けながら各グループの動向をまとめる。

グループ1では、セキュリティや災害の分野との関わりが強いことがうかがえる^{9,16)}。理由として、IoT の製品が現実生活に浸透するにつれ、それに伴いセキュリティ脅威の標的になるデバイスが増えることにつながるからである。実際に、現実世界で車載の組み込み OS がハッキングされ、小売業などで販売管理に用いられる「POS 端末」がウイルス感染した事例が発生している。今後、IoT 関連の製品が爆発的に増える一方、それに伴いセキュリティの確保を急がなくてはならない。

グループ2では、医療・福祉業界との関わりが強いことがうかがえる¹⁰⁾。理由として、医療機器に通信機能を持たせることにより、遠隔でも患者の健康状態を把握することや、一人一人の健康意識の向上も見込まれる。例えば、ユーザーの虫歯予防

に役立つ「コネクテッド歯ブラシ」や「GlowCap」という薬服用管理デバイスなど既に市場に現れている。また、携帯機器と組み合わせることにより、既存の医療行為を支援するようなソフトが注目を浴びている。実際に、健康データを計測・収集するシステムが開発されている。このことから、医療分野における IoT は、今後更なる発展を遂げていくと考えられる。

グループ3では、住宅・建設業界との関わりが強いことがうかがえる¹¹⁾。住宅に IoT に関連した設備を搭載することにより、より快適な暮らしをサポートできると予想できる。住環境において、人の居場所を検知することで、犯罪抑止力や家電製品の遠隔操作によるオン・オフなど、新たなサービスを提供できるのではないかと考えられる。

この分野の動向として特に注目するものとして、大手住宅メーカーが東京大学と共同研究した成果が挙げられる。温度・湿度などを検知する「環境モニタリングシステム」や自分の居場所を自動認識する「ユビキタス場所情報インフラ」などが研究用に大学に提供されている。このような研究に象徴されるように、住環境において新たなサービスが検討されており、近い将来に IoT を用いた住宅が身近になると考えられる。

グループ4では、産業・観光分野との関わりが強いことがうかがえる¹²⁾。理由として、ネットワーク接続されたセンサーやソフトウェアにより、複雑で高度な機器や設備を統合することで、様々な産業分野においてイノベーションを起こす可能性があると考えられる。その概念が Industrial Internet (産業機器と IT の融合) である。

グループ5では、ものづくり業界、企業の業務システムとの関係である¹³⁾。IoT は、生産の現場や企業活動の場でも威力を発揮するはずである。現在、シリコンバレーにおいて IoT のものづくりは盛んに行われているようである。

図表4 IoT 研究の6つのグループとその対象領域

	IDF値が高いキーワード	対象領域
グループ1	information security, european disasters	セキュリティ, 災害の分野
グループ2	healthcare	医療・福祉業界
グループ3	building automation systems	住宅・建設業界
グループ4	wisdom scenic, industrial informatics	産業分野, 観光分野
グループ5	enterprise systems, device collaboration	ものづくり業界, 企業の業務システム
グループ6	benchmarking, gamification	IT・コンピュータ業界, サービス業界

グループ6は、IT・コンピュータ業界、サービス業界との関わりが強いことがうかがえる¹⁴⁾。このグループの論文からうかがえることとして、将来多くのIoT製品・サービスが開発されるが、その製品を制御するサービスが多種多様に開発される時代になる。また、IoT製品に適用されるセキュリティパッチやウイルスソフトなど、関連分野で多くの新しいサービスが必要となるだろう。

6 まとめ

学術論文データベースが整備され、分析のためのツールもオープンソースソフトウェアの形で入手できる¹⁷⁾。本稿で述べたような手法が比較的容易に実践でき、計量書誌学的なアプローチによる研究動向調査の可能性が広がっている。今回の調査結果の中では、IoTについての学術論文の発表数が2009年

度から急激に増大したことは注目に値する。このことから、IoTが現在ホットな研究対象であることは明確である。また、IoTと関係性の深いキーワードをTF-IDFによる単語の頻度分析に基づいて考察すると、ネットワーク関連と通信技術の分野との関わり合いが特に強いことがうかがえる。また、IoTという単語は、ユビキタスとM2Mの両方の分野を兼ね合わせたものであり、関連する通信技術分野についてのキーワードが多く収集された。

2014年に実施された第10回科学技術予測調査においても、ICT分野で重要度が上位となった課題には、「エクサ～ゼタバイトスケールのHPC・ビッグデータ処理技術の社会現象・科学・先進的ものづくりなどへの適用による革新」と「介護・医療の現場で、患者の状態をリアルタイムに把握し、その状態に最適なケアを低コストで提供するシステム」が挙げられている¹⁵⁾。IoTが今後こうした分野の核となることは疑いなく、今後も動向を注視することが重要である。

参考文献

- 1) 文献データベース“IEEE Digital Library Xplore”、法政大学図書館電子ジャーナル
- 2) L. ライデスドルフ、「科学計量学の挑戦」玉川大学出版部、2001年
- 3) Toby Segaran「集合知プログラミング」オライリー・ジャパン
- 4) MTmineR：<https://code.google.com/p/mtminer/>
- 5) 野村総合研究所 城田真琴、「Internet of Things (モノのインターネット) 時代の到来」、2012年11月27日：http://www.nri.com/jp/event/mediaforum/2012/pdf/forum183_1.pdf
- 6) 国土交通省「社会インフラのモニタリング技術活用推進検討委員会」資料：<http://www.mlit.go.jp/tec/monitoring.html>
- 7) 総務省「2020年代に向けたモバイル分野の競争政策」：http://www.soumu.go.jp/main_content/000313742.pdf
- 8) 経済産業省産業構造審議会商務流通情報分科会資料「IoT時代に対応したデータ経営2.0の促進」：http://www.meti.go.jp/committee/gizi_1/32.html
- 9) Eleana Asimakopoulou, Nik Bessis^{2,1}, Stelios Sotiriadis², Fatos Xhafa³ and Leonard Barolli “A Collective Intelligence Resource Management Dynamic Approach for Disaster Management: A Density Survey of Disasters Occurrence” 2011 Third International Conference on Intelligent Networking and Collaborative Systems
- 10) Mohammadreza S. Shahamabadi, et.al “A Network Mobility Solution Based on 6LoWPAN Hospital Wireless SensorNetwork (NEMO-HWSN)” 2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing
- 11) Dhananjay Singh, et.al, “A survey of Internet-of-Things: Future Vision, Architecture, Challenges and Services”, 2014 IEEE World Forum on Internet of Things (WF-IoT)
- 12) Zhenfeng SHAO, Chong LIU, “Intelligent management and service for Wisdom Scenic Based on Internet of Things” Proceedings of the 2011 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems
- 13) Wu He and Li Da Xu, “Integration of Distributed Enterprise Applications: A Survey”, IEEE Trans. on Industrial Informatics, Vol. 10, No. 1, Feb. 2014

- 14) Internet of Things for Smart Cities, IEEE Internet of Things Journal Vol.1 No.1, Feb. 2014
- 15) 第10回科学技術予測調査結果速報 全体概要、科学技術・学術政策研究所、2014年11月：
<http://www.nistep.go.jp/archives/18742>
- 16) 藤井章博、「ユビキタスネット社会のコンテクストアウェアネス技術研究の動向と課題」、科学技術動向、2007年8月、No.77、p.18-25：<http://hdl.handle.net/11035/1867>
- 17) 藤井章博、「技術文書に見るインターネット要素技術の動向」、科学技術動向、2014年1月、No.142、p.19-24：
<http://hdl.handle.net/11035/2474>

..... **執筆者プロフィール**



藤井 章博

科学技術動向研究センター 客員研究官

博士(工学)。法政大学理工学部教授。分散コンピューティングと通信プロトコルの研究に従事した後、電子商取引システムの構築プロジェクトを実施。現在、情報通信技術のイノベーションが経営や政策に与える影響に興味を持つ。