

SIGIR2018 参加報告

Report on SIGIR2018

野本 昌子^{1*} 富樫 陸¹ 鈴木 翔吾¹ 椎野 弘章¹
Masako Nomoto¹ Riku Togashi¹ Shogo Suzuki¹ Hiroaki Shiino¹

¹ ヤフー株式会社

¹ Yahoo Japan Corporation

Abstract: The 41st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018) took place in Ann Arbor, USA in July 2018. A total of 409 papers for the full paper track were submitted and reviewed by six tracks, for an acceptance rate of 21%. About 46.0% of the accepted papers are related to neural network, distributed over about 73.1% of technical sessions.

1 はじめに

情報検索 (Information Retrieval, 以下 IR) のトップカンファレンス SIGIR2018 が 2018 年 7 月に米国で開催され、約 740 名が参加した¹.

本稿では SIGIR2018 の参加報告として、SIGIR の概要、SIGIR2018 の概要と動向、論文の一部、次回の概要と国内のコミュニティの動向を紹介する。

2 SIGIR

SIGIR (SIGIR Conference on Research and Development in Information Retrieval) は IR 分野のトップカンファレンスで、主催組織は検索・情報アクセス分野の国際学会 ACM の IR の分科会 SIGIR(Special Interest Group on Information Retrieval)²である。SIGIR は 1978 年から毎年開催されており、ACM SIGIR が後援している国際会議の中ではもっとも歴史が古い。

IR の周辺分野には自然言語処理、人工知能、データマイニング、Human Computer Interaction などがある([1])。

Microsoft Academic の報告 [2] によると、SIGIR の論文は CIKM³, WWW⁴, WSDM⁵等、周辺分野の国際会議と相互に論文の参照関係がある。2015 年から 2017

年の間に SIGIR の論文を参照した venue, SIGIR の論文が参照した venue を各々、図 1⁶, 図 2⁷に示す。SIGIR の論文は上に述べた 3 会議の他に、理論寄りの ICTIR⁸, 情報推薦の RecSys⁹, データマイニング系の KDD¹⁰等から参照されている。逆に SIGIR の論文からも上述の 3 会議に加えて、KDD, 機械学習系の NIPS¹¹や ICML¹², 自然言語処理の ACL¹³や EMNLP¹⁴, 人とコンピュータのインタラクション (HCI) に関する CHI¹⁵など、多様な会議の論文を参照している。

図 3 に SIGIR の大陸別開催回数を示す¹⁶。従来 SIGIR は欧米で開催されていたが、近年はアジアとオセアニア、南北アメリカ、ヨーロッパを順に回っている。

3 SIGIR2018

3.1 開催概要

SIGIR2018 は米国ミシガン州アナーバーのミシガン大学で、7 月 8 日から 12 日まで 5 日間に渡り開催され

⁶[2] のグラフ Top Venue Citation Over Time の 2015 年から 2017 年のデータを元に作成。

⁷[2] のグラフ Top Venue Reference Over Time の 2015 年から 2017 年のデータを元に作成。

⁸ACM International Conference on the Theory of Information Retrieval

⁹ACM Conference on Recommender Systems

¹⁰ACM SIGKDD Conference on Knowledge Discovery and Data Mining

¹¹Conference on Neural Information Processing Systems

¹²International Conference on Machine Learning

¹³Annual Meetings of the Association for Computational Linguistics

¹⁴Empirical Methods in Natural Language Processing

¹⁵ACM CHI Conference on Human Factors in Computing Systems

¹⁶[3] を元に作成。

*連絡先: ヤフー株式会社

千代田区紀尾井町 1-3

E-mail: mnomoto@yahoo-corp.jp

¹SIGIR2018 のオープニングの紹介によると、登録者数 740 名、うち約 40%は企業と非アカデミアの組織から参加。

²本稿では区別のため、以降、'ACM SIGIR' と記す。

³ACM International Conference on Information and Knowledge Management

⁴ACM International World Wide Web Conference

⁵ACM International Conference on Web Search and Web Data Mining

表 1: SIGIR2018-2016 の概要

会議 (回)	SIGIR2018(41)	SIGIR2017(40)	SIGIR2016(39)
会期	July 8-12	Aug 7-11	July 17-21
開催地	Ann Arbor Michigan, USA	Tokyo, Japan	Pisa, Italy
参加者数	740	911	566
フルペーパー	86/409 (21%)	78	62
ショートペーパー	98/327 (30%)	121	104
デモ	18/36 (50%)	17	21
チュートリアル	11/13 (85%)	8	12
ワークショップ	10/15 (67%)	8	7
SIRIP(Industry Track)	9/16 (56%)	12	12

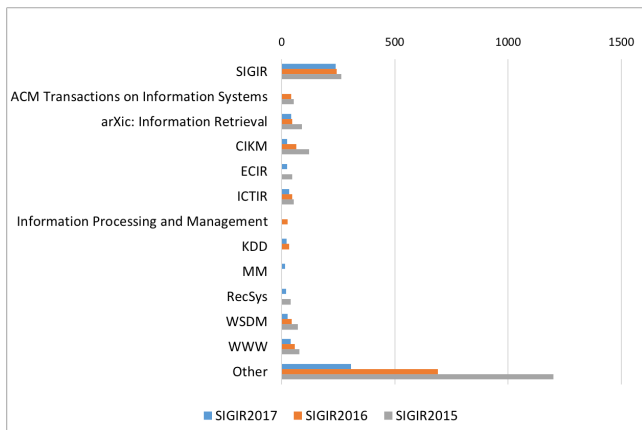


図 1: SIGIR の論文を参照した論文の venue (SIGIR2015-SIGIR2017)

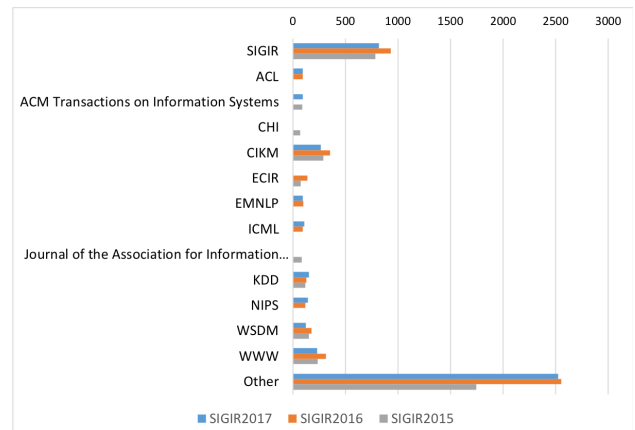


図 2: SIGIR の論文が参照した論文の venue (SIGIR2015-SIGIR2017)

た. 表 1¹⁷に過去 3 回の SIGIR の開催概要を示す. 今回のプログラムも例年通りチュートリアルとワークショップが各 1 日, 本会議が 3 日間であった. 大きな変化は, 企業の IR の応用に関する SIRIP¹⁸が 2 日間に拡大されたことである.

過去最大の 911 名の参加者を集めた昨年に続き, 今回も約 740 名 (登録者数) と多数の参加があった. なお, 2001 年以降で SIGIR の参加者数が 700 名を超えたのは 2006 年以来である ([4].p.9).

日本からは, アカデミアから少なくとも 13 名, 企業からはヤフー, リクルート住まいカンパニー, リクルートジョブズ, メルカリ, NTT コムウェア, Honda R&D 等から少なくとも 17 名が参加した.

表 2 に今回のスポンサーを示し, 前回, 前々回のスポンサーに該当する場合は'*' を付した¹⁹. スポンサー全体としては米国と中国の企業が多い. 前回は開催国日本の企業が目立っていたが, 今回は見られない. 情報

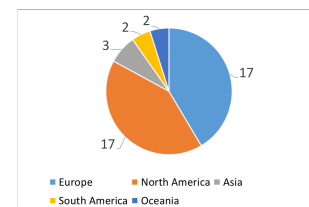


図 3: 大陸別開催回数

通信の企業が多いのは例年通りで, 米国企業の顔ぶれは変わらず, アジア企業が増加を続けている. 今回スポンサーに加わったのは中国の車のライドシェアサービスの DiDi, 音楽ストリーミングサービスの Spotify, 米国の小売業や e コマースの企業である.

3.2 投稿・採択状況

図 4 に SIGIR のフルペーパーの投稿数, 採択数, 採択率の推移 (1999 年から 2018 年まで) を示す²⁰. 今回は昨年を上回る 409 本のフルペーパーが投稿され, うち 86 本が採択され, 採択率は 21.0%であった. SIGIR

¹⁷フルペーパー以降の行は SIGIR2018 については採択数/投稿数 (採択率), 他は採択数を示す.

¹⁸The SIGIR Symposium on IR in Practice. Industry Track に相当する.

¹⁹[5]pp.xxxviii-xxxix, [6]pp.xxxv-xxxvi, [7]pp.xvii-xviii を元に作成.

²⁰[5] の Publication, Acceptance Rate の表より作成.

表 2: SIGIR2018 のスポンサー

国	分類	企業	2017	2016
アメリカ	情報通信	Microsoft	*	*
		IBM	*	*
	小売/e コマース	Bloomberg		*
		Google	*	*
		ebay	*	*
		amazon	*	*
		wayfair		
		Home Depot		
中国	情報通信	Alibaba	*	
		Baidu	*	*
		Sogou	*	
	旅行 (ライドシェア)	Byte Dance		
		kika		
		DiDi		
韓国	情報通信	NAVER	*	
スウェーデン	音楽	Spotify		

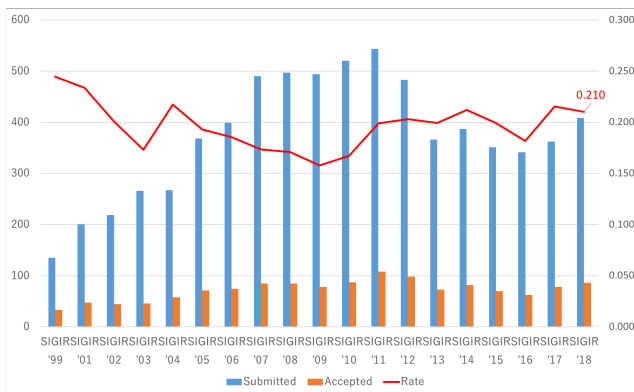


図 4: フルペーパーの投稿数, 採択数, 採択率の推移 (1999 年から 2018 年まで)

全体 (1999 年から 2018 年) の平均採択率 19.1% に比べると今回はやや易しく, 昨年の 21.5% と同程度である。

全投稿論文の国別著者数 (2016 年から 2018 年) を図 5²¹ に示す。2016 年時点の投稿数は 1 位が米国, 2 位が中国であったが, 2017 年に 1 位と 2 位が入れ替わった。

次に, フルペーパーの採択論文の国別著者数の割合 (2016 年, 2018 年) を図 6²² に示す。2016 年の時点では米国が 34%, 中国が 23% であり, 昨年も順位の変動はなかったが²³, 今回はついに中国が 34%, 米国が 30% と順位が入れ替わった。前回のフルペーパーは日本から少なくとも 3 本が採択されたが, 今回は日本からの採択はなく, 本会議の Interface のセッションでは梅本ら [9] の TOIS (ACM Transactions on Information Systems) のジャーナルの論文の口頭発表²⁴があった。

著者の所属の分布を見るため, SIGIR2018 の全著者の所属をアカデミア, 企業, 両方, それ以外に分類し

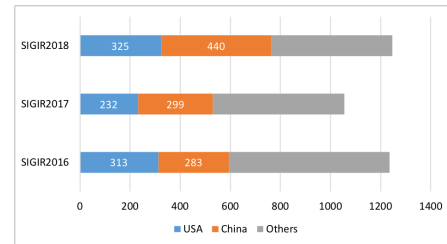


図 5: 全投稿論文の国別著者数 (2016 年-2018 年)

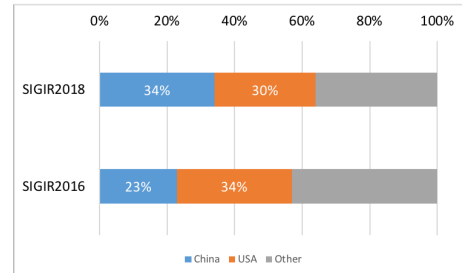


図 6: フルペーパー採択論文の国別著者数の割合 (2016 年, 2018 年)

たところ, SIGIR2018 の全著者の約 78% がアカデミア, 約 20% が企業の著者であった²⁵。

次に, SIGIR2016 から SIGIR2018 のフルペーパーの採択論文の著者の所属を同様に分類した結果を図 7 に示す²⁶。アカデミアの著者が年々増えているが, 企業の著者数は年により異なる。SIGIR2018 ではアカデミアが 81.6%, 企業が 15.6% であった。

さらに, 2016 年から 2018 年のフルペーパーの企業別著者数を図 8 に示す²⁷。2016 年は Yahoo と Microsoft が突出していたが, 2017 年以降は Microsoft が単独で他を引き離している。今回は特に Spotify 等, 新しいスポンサーが目立つ。

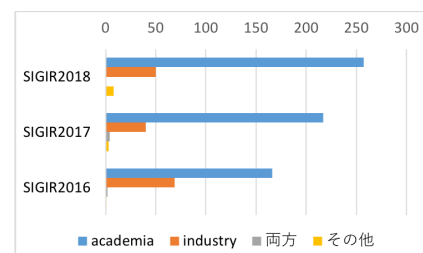


図 7: フルペーパーの著者の所属グループ別著者数 (2016 年から 2018 年)

²⁵ [11] の full papers, short papers, demo papers, SIRIP Industry Papers, Doctoral Consortium のいずれかの表に含まれる著者の所属を元に分類した。

²⁶ [11] の full paper の著者の所属, [12], [13] のプログラムの著者の所属を元に作成。

²⁷ [11] の full paper の著者の所属, [12], [13] のプログラムの著者の所属を元に作成。

²¹ [5] p.iii, [6] p.iii, [7] p.iii を元に作成。

²² [5] p.iii, [7] p.v を元に作成。

²³ [8], SUBMISSIONS BY COUNTRY (p.6)。

²⁴ SIGIR2017 から TOIS の論文の著者は CHIIR, SIGIR, IC-TIR のいずれかの国際会議で口頭発表を行えることになった [10]。

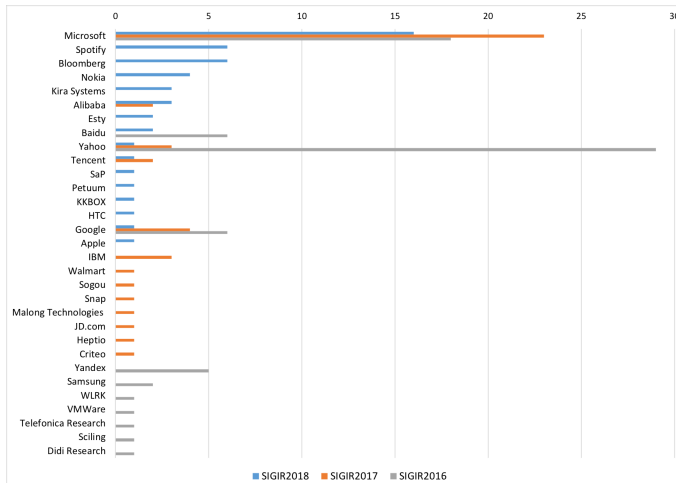


図 8: フルペーパーの企業の著者の所属 (2016 年から 2018 年)

表 3: SIGIR2018 のトラック

トラック
(a) Search and Ranking
(b) Foundations and Future Directions
(c) Domain-Specific Applications
(d) Content Recommendation, Analysis and Classification
(e) Artificial Intelligence, Semantics, and Dialog
(f) Human Factors and Interfaces

3.3 トラック

今回は例年と異なり、表 3 に示した 6 つのトラック別に投稿、採択が行われた。投稿の多いトラックは (d) のコンテンツ推薦で、124 件 (投稿論文の 30.0%) の投稿があった²⁸。参考のため、前回の推薦に関する投稿を図 9 のトピック別投稿数で見ると、推薦システム (recommender system) が 50 件以上 (投稿論文の 13.8% 以上)、推薦のパーソナライズ (personalized recommendation) が 40 件以上 (同 11.0% 以上)、推薦アルゴリズム (recommendation algorithms) が 40 件弱 (同 11.0% 弱) であった。推薦に関する投稿が多い傾向自体は前回から続いている。

一方、今回採択数の多いトラックは (c) の分野に特化した応用と (f) Human Factors で各々 20 件 (採択論文の 23.2%) である ([5]p.iii)。参考のため、前回の採択論文の中で多かったトピック²⁹を見ると、行動分析 (15%) やユーザスタディ (14%) が含まれている。採択論文にユーザに関する研究が目立つのも昨年から続く傾向と考えられる。

²⁸[5]p.iii.

²⁹The most popular call-for-papers topics for accepted papers were ranking algorithms (19%), text representation (17%), behavioral analysis (15%), recommender systems (14%), and user studies (14%). ([6]p.iii)

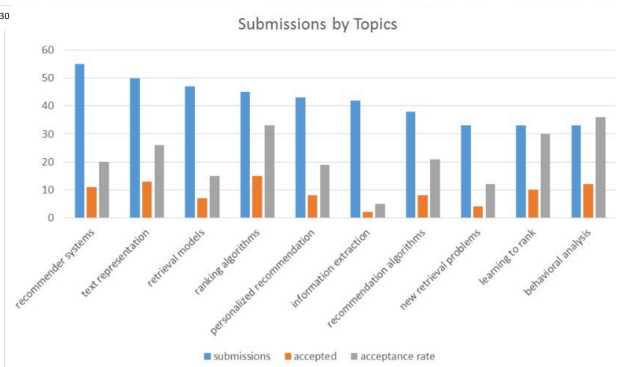


図 9: トピック別投稿数 (SIGIR2017, 投稿数順)

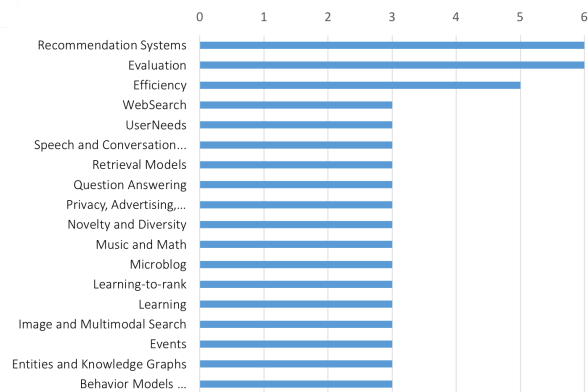


図 10: セッション別論文数 (SIGIR2016)

3.4 セッション

SIGIR2016, SIGIR2017, SIGIR2018 のセッション別の論文数を図 10, 図 11, 図 12 の青線で示す。

2016 年は推薦, 評価, 効率に関するものが多く、新しいセッションでは音声と対話やセマンティック・ウェブのエンティティと knowledge graph があった。

2017 年はフィルタリングと推薦、検索モデルとランキング、検索のインタラクションなどが多かった。また、2016 年に続き、エンティティや、対話と質問応答のセッションもあった。

今回はランキング学習の論文数が突出しているように見えるが、ユーザーの行動の分析に関するもの、推薦に関するものも複数セッションにまたがっている。対話やセマンティック・ウェブ関連のセッションは定着した感がある。新しいセッションの一つにソーシャルグッド (Social Good) があり、これは基調講演のキーワードでもある。

3.5 ニューラルネットに関する動向

SIGIR2016 の基調講演で Christopher Manning 氏 (スタンフォード大学) は今後 2, 3 年は深層学習が SI-

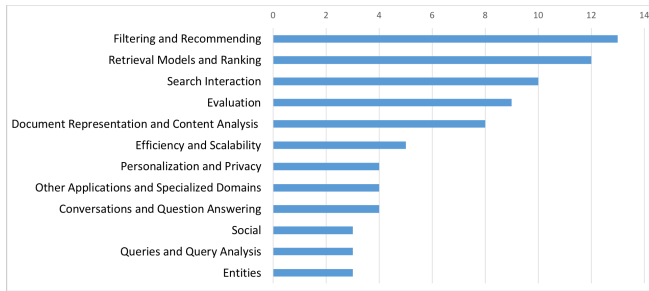


図 11: セッション別論文数 (SIGIR2017)

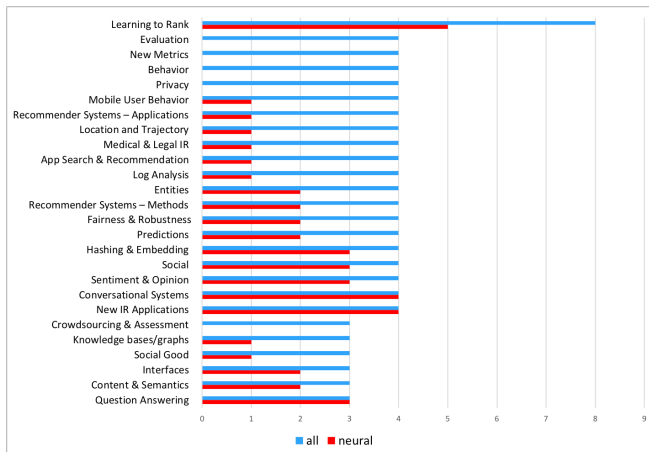


図 12: セッション別の論文数及びニューラルネット関連論文数 (SIGIR2018)

GIR を支配すると予測した [14]. また, 前回 SIGIR2017 の最大の話題はニューラルネットで, Neural Information Retrieval のワークショップ (Neu-IR)³⁰ のオープニングでは約 4 分の 1 の論文が neural IR に関するものであり, 「ニューラルネットの年 (The year of the Neural Nets)」と報告された.

そこで, SIGIR2018 の本会議のテクニカルセッションの論文を対象にニューラルネットに関する論文かどうかを調べたところ³¹, フルペーパーの 40/87 本 (約 46.0%), TOIS の論文も含めると 45/102 本 (約 44.1%) が該当した. 各セッションの発表論文のうち, ニューラルネット関連の論文数を図 12 の赤線で示す. さらにセッション単位の分布を見ると, フルペーパーを対象とするとテクニカルセッションの約 73.1%(19/26)³², TOIS の論文も含めると同約 80.8%(21/26) でニューラルネット関連の論文が見られる.

³⁰<https://neu-ir.weebly.com>

³¹調査は基本的にはキーワードベース ('neural net', 'deep learning') で, future work として挙げたもの等, 非該当と容易に判別できるもののみ除いた.

³²Session 1D: Learning to Rank I と 3D: Learning to Rank II はまとめて 1 セッションとカウントした.

3.6 基調講演

シカゴ大の研究助教³³の Ghani 氏はかつてのオバマ大統領の選挙キャンペーン (2012 年) のデータサイエンスチームの中心人物で, シカゴ大ではデータサイエンスティストが実社会の問題解決に取り組むようにするための活動³⁴に関わっている. 講演 [15] ではソーシャルグッド³⁵と公共政策を取り上げ, 政策や社会問題の解決に機械学習やデータサイエンスを役立てる取組み等を紹介した.

Gerald Salton 賞³⁶を受賞したタンペレ大学教授の Jarvelin 氏は Information Seeking や IR の評価に関する研究で著名で, IR の適合性の評価指標 nDCG の最初のバージョンを提案した ([16], pp.32-33) 人物である. 講演 [17] は情報のインタラクションに関するもので, これまでの研究を振り返り, 情報科学の研究は, 人間を仕事や生活の文脈で情報の環境を認識してそれらとインタラクションするアクターとして研究するものと考えたこと, IR の研究でも人に関わる部分が重要であることを強調した.

3.7 チュートリアル

今回は表 4 に示す 11 のチュートリアルがあった³⁷.

(f) の深層学習のチュートリアル [18] では, 検索におけるマッチングと推薦におけるマッチングを取り上げ, それらは各々 SIGIR, RecSys³⁸という異なるコミュニティで研究されているが, 統一的な視点で整理できることを示した上で, さらに, マッチングの伝統的な手法から最新の深層学習の手法までを概観した.

3.8 SIRIP(Industry days)

SIRIP³⁹は, 企業での IR の応用に関するもので, SIGIR2016, SIGIR2017 では本会議中の 1 日に, テクニカルセッションと並行して行われていたが, 今回は 2 日間 (Industry Days) で, 基調講演, 招待講演と一般の講演があった.

アマゾンの Carmel ら [19] からはアレクサショッピングの商品に関する QA(質問応答)の研究とその課題等の紹介があった. 研究では商品に関する主観的な質問にフォーカスしているとし, 例えば, 贈り物のアイデアへの意見や具体的な商品のアドバイス, 商品の

³³Research Assistant Professor.

³⁴Data Science for Social Good Summer Fellowship

³⁵社会に対して良いインパクトを与えようとする活動やサービス, 製品などのこと.

³⁶Gerald Salton 賞は IR の研究に継続的に顕著な貢献を行ってきた人に対して 3 年毎に贈られる賞.

³⁷<http://sigir.org/sigir2018/program/tutorials/>

³⁸ACM Conference on Recommender Systems

³⁹<https://sigir.org/sigir2018/program/industry-days/>

表 4: Tutorial(SIGIR2018)

Tutorial	説明
(a) Knowledge Extraction and Inference from Text:	知識抽出・推論
(b) Probabilistic Topic Models for Text Data Retrieval and Analysis	テキストデータ検索・分析向け確率的トピックモデル
(c) SIGIR 2018 Tutorial on Health Search (HS2018)	医療・健康に関する検索
(d) Conducting Laboratory Experiments Properly with Statistical Tools	統計的ツールを用いる (実験室) 実験の方法
(e) Information Discovery in E-commerce	E コマースにおける情報発見
(f) Deep Learning for Matching in Search and Recommendation	深層学習の検索と推薦におけるマッチングへの応用
(g) Generative Adversarial Nets for Information Retrieval	敵対的生成ネットワーク (GAN) を用いた IR
(h) Tutorial on Utilizing Knowledge Graphs for Text-centric IR	Knowledge Graph を用いたテキスト情報検索
(i) Neural Approaches to Conversational AI	ニューラルネットを用いた対話型 AI
(j) Efficient Query Processing Infrastructures	効率的なクエリ処理の基盤
(k) Fusion in Information Retrieval	IR におけるフュージョン

使い方, 特定商品の比較や商品群の中でのおすすめへのコメント等を例に挙げた. 回答にはカタログや web データの他, レビューや QA などにも利用している. 商品検索の研究課題として音声インターフェイスのフレンドリーな回答生成, 分散した異種の情報からの回答生成, 個人適応等の他, 特に主観的な質問に関する課題としては複数の観点からの回答 (multi-aspect answer), 質問の評価を挙げた.

マイクロソフトの Agrawal ら [20] はインド市場向けのチャットボット, ルー (Ruuh) の開発経験を紹介した. インド市場は若年層が多く, ルーは全ての若者の BFF(永遠の大親友) となることを目指しており, 単に適合した応答を返すだけでは不十分である. 検索ベースのチャットボットで対話と応答のペアをインデクシングして会話に再利用しているが, twitter 等の生データのクレンジングでは攻撃的な発言や個人を特定できる情報を除くほか, 実際にはできないことを言う表現 (例: 電話するね, 等) も削除している. ユーザのシェアした画像や動画へのコメントを求められることも多く, 感情や意見の表出が期待されている. また, インドでは言語が多様で会話の 3-5% でコードミクシング (複数言語の混ざった発話) があり, ボットがコードミクシングで応答すると利用可能性は 30% にまで上がるという.

リンクトインの Geyik ら [21] からは優秀な人材を検索・推薦するシステム LinkedIn Recruiter の紹介があった.

3.9 ワークショップ

今回は表 5 に示す 10 のワークショップがあった⁴⁰. 全日のワークショップは検索への対話的アプローチ, knowledge graph の利用, e コマースと前回からの継続が多い. 新しいワークショップとしては特定分野の専門家による複雑な検索の他, ITS に関するもの, 驚きの定量化, 理由を説明可能な推薦と検索等があった.

e コマースのワークショップでは商品タイトルから商品のカテゴリを予測する機械学習のデータチャレンジ

(Rakuten Data Challenge⁴¹) の結果が報告された. カテゴリは ID の木構造 (leaf は 3,000 強) である. 商品タイトルとカテゴリ ID のパスのペア (0.8M) で学習して, テスト用の商品タイトル (0.2M) のカテゴリを推定するタスクで, 末端まで推定できた場合のみ正解とする. 1 位の Skinner[22] の手法は character ベースの LSTM と F1 最適化を組み合わせたものであった.

3.10 SIGIR2018 のトレンド

今回のトレンドを以下にまとめる.

- ニューラルネットに関する論文が大幅に増加 (フルペーパーの約 46.0%, テクニカルセッションの 73.1% に分布)
- 投稿数, 採択数とも中国からの論文が増加
- トラック別投稿数はコンテンツの推薦が多く, 前回に続き, 推薦に関する論文の投稿が多い
- トラック別採択数は分野特有の応用と Human Factors が多く, 今回もユーザよりの論文が多い
- 新しいキーワードはソーシャルグッド (social good)
- 新しいワークショップは分野の専門家の複雑な検索, ITS 関連, 驚きの定量化, 理由を説明可能な推薦と検索等
- 対話やセマンティック・ウェブのセッションは定着

3.11 テクニカルセッション

テクニカルセッションのうち, 一部の論文を紹介する.

3.11.1 受賞論文

Best paper はフェアネスとロバストネスのセッションの Canameres ら [23] で, 推薦システムにおける人気 (popularity) の扱いに関するものである. 近年, 推薦システムの評価やアルゴリズムでは人気が重視され

⁴⁰<http://sigir.org/sigir2018/program/workshops/>

⁴¹<https://sigir-ecom.github.io/data-task.html>

表 5: Workshop (SIGIR2018)

Workshop	回	区分	トピック
CAIR ' 18	2	全日	検索への対話的アプローチ. マルチターンの複雑な対話の支援, マルチモーダルインターフェイス.
ECOM ' 18	2	全日	e コマース (商品検索と検索).
ProfS2018	1	全日	特定分野の専門家による複雑な検索.
KG4IR ' 18	2	全日	検索での knowledge graph 利用.
BIRNDL ' 18	3	半日	自然言語処理, IR, サイアントメトリクス, テキストマイニング, 推薦による学術文献の理解/分析/検索.
CompS ' 18	1	半日	Computational surprise(驚きの定量化, どのようにして IR に取り込むか, 等).
DATA:SEARCH ' 18	1	半日	構造化/半構造化されたデータの検索.
EARS ' 18	1	半日	理由を説明可能な推薦と検索.
Intelligent Transportation Informatics	1	半日	道路交通に関する異種情報の獲得, 統合, 分析.
Learning from Limited/Noisy data for IR	1	半日	IR のための限られたデータまたはノイズの多いデータからの学習.

る一方、バイアスとして取り除くべきとする研究もあり、答えが出ていない。そこで、推薦に人気を用いる効果、オフラインの実験で効果を正確に測定できているかなどを調べ、結論としては人気に基づく推薦が良いかどうかは一概には決められないことを示した。

3.11.2 ログ分析

Gathright ら [24] は音楽推薦システムの精度向上のため、ユーザの満足度に関する詳細な調査と分析を実施し、満足度の予測を行った。Spotify の音楽推薦を 10 週間利用したユーザを対象に利用目的やユーザ体験に関するインタビューを行い、さらにユーザ約 18,500 人を対象に利用目的に加えて週毎と全体的な満足度のサーベイを実施した。これらの分析の結果、音楽推薦システムの 4 つの利用目的を抽出し、利用目的によってユーザ行動が異なること、ポジティブなユーザ体験が具体的にどのような行動に現れるか等を明らかにし、満足度の予測モデルを構築した。さらにユーザの評価におけるバイアスやユーザが満足するための条件、満足度の推定に役立つもの等の知見も得た。

3.11.3 行動

Lu ら [25] はモバイル環境でニュース記事を読む際のユーザの preference⁴²の変化や、その変化に影響を与えるものをユーザスタディをふまえて分析した。ユーザの preference は記事を読む前後とタスクの後の 3 段階を区別し、タスクが終わった後の preference を真の preference としている。得られた知見の例としては、クリックはユーザの preference の指標としては不十分であること、記事への preference は記事を読む前後、タスク後の 3 段階で変化すること、preference の変化に影響を与えるのは記事の品質、ユーザインタラクションのコンテキストで、例えば直前に読んだ記事の真の preference が低いと次の記事を過度に高く評価する傾向がある、等である。さらにユーザの行動、記事の品

質、ユーザのインタラクションのコンテキストも用いてユーザがクリックした記事の真の preference を予測するモデルを構築し、従来のクリックベースの指標を予測した preference で置き換えることでユーザの満足度を改善できることを示した。

4 SIGIR2019

次回, SIGIR2019 は第 42 回で, 2019 年 7 月 21 日から 25 日までパリ (会場: シテ科学産業博物館) で ARIA⁴³の後援を受けて開催される。すでにフルペーパーの CFP, 会場, オーガナイザの情報は公式ページ⁴⁴に掲載されており, 今後の更新情報については, 同ページに加えて公式 Twitter⁴⁵を参照されたい。

5 国内のコミュニティの動向

日本で開催した SIGIR2017 の成功を受けて, 昨年 12 月, 主催組織の ACM SIGIR の初めての支部である東京支部⁴⁶が誕生した。SIGIR 関連のセミナー等のイベントや会員向けの情報発信を行い, 国内外の研究コミュニティの交流を促進し, IR 研究における日本のプレゼンスの向上を目指す。2018 年 10 月には今回の SIGIR2018 を含む最近の主要な国際会議の論文 (約 27 本) の読み会 (IR Reading2018 秋)⁴⁷を開催した。

6 まとめ

SIGIR は IR 分野の最高峰の国際会議で, 今回の SIGIR2018 は米国で開催され, 約 740 名が参加した。フルペーパーの採択率は約 21 % (86/409) で, 国別に見ると投稿数, 採択数とも中国の躍進が目立った。トレ

⁴²本論文では 'like' を 'preference' と呼んでいる。

⁴³French Association for Information Retrieval and Applications

⁴⁴<http://sigir.org/sigir2019/>

⁴⁵<https://twitter.com/sigir2019>

⁴⁶Tokyo ACM SIGIR Chapter, <http://sigir.jp/>

⁴⁷http://sigir.jp/post/irreading_2018fall/

ンドはニューラルネット (フルペーパーの約 46.0%, テクニカルセッションの約 73.1%に分布), トラック別投稿数ではコンテンツ推薦が多く, 前回に続き推薦に関する投稿が多い. トラック別採択数では分野に特化した応用と Human Factors に関する論文が多く, ユーザよりの論文が多い傾向が続いている. その他のキーワードとしてはソーシャルグッドがあった.

次回 SIGIR2019 は 2019 年 7 月にパリで開催される. 国内のコミュニティの動向としては, 2017 年 12 月に ACM SIGIR 初の地方支部である東京支部が誕生した. 今後の SIGIR において, 日本のプレゼンスがさらに向上することを願っている.

参考文献

- [1] Toshihiro Kamishima: ML, DM, and AI Conference Map, <http://www.kamishima.net/archive/MLDMAImap.pdf>. (2018)
- [2] Microsoft Academic: SIGIR Conference Analytics, <https://www.microsoft.com/en-us/research/project/academic/articles/sigir-conference-analytics/>. (2018)
- [3] SIGIR: List of SIGIR Locations and History, <http://sigir.org/general-information/history/>.
- [4] 樗 惇志: SIGIR2017 参加 (開催) 報告-会議概要とニューラルネットな情報検索-, 第 11 回テキストアナリティクス・シンポジウム, <http://www.ieice.org/~nlc/attachment/NLC20170908-SIGIR-2017-Keyaki.pdf>. (2017)
- [5] The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, <https://dl.acm.org/citation.cfm?id=3209978>. (2018)
- [6] Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, <https://dl.acm.org/citation.cfm?id=3077136>. (2017)
- [7] Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, <https://dl.acm.org/citation.cfm?id=2911451>. (2016)
- [8] Arjen de Vries, Hang Li, Ryen White: ACM SIGIR 2017 - Opening - PC Chairs, <https://www.slideshare.net/arjenpdevries/acm-sigir-2017-opening-pc-chairs>. (2017)
- [9] Kazutoshi Umemoto et al.: Search by Screenshots for Universal Article Clipping in Mobile Apps, ACM Transactions on Information Systems (TOIS) - Special issue: Search, Mining and their Applications on Mobile Devices, Vol.35, Issue.4, pp.34:1-2. (2017)
- [10] SIGIR: TOIS Presentation, <http://sigir.org/conferences/tois-presentation/>. (2017)
- [11] ACM SIGIR 2018 Conference Organization: ACM SIGIR 2018 Accepted Papers, <http://sigir.org/sigir2018/accepted-papers/>. (2018)
- [12] SIGIR2017, <http://sigir.org/sigir2017/>. (2017)
- [13] SIGIR2016, <http://sigir.org/sigir2016/>. (2016)
- [14] Christopher Manning: Natural Language Inference, Reading Comprehension and Deep Learning, <https://nlp.stanford.edu/manning/talks/SIGIR2016-Deep-Learning-NLI.pdf>, p.72. (2016)
- [15] Rayid Ghani: Data Science for Social Good and Public Policy: Examples, Opportunities, and Challenges, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, p.3. (2018)
- [16] 酒井哲也: 情報アクセス評価方法論, コロナ社. (2015)
- [17] Kalervo P. Jarvelin: Salton Award Keynote: Information Interaction in Context, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp.1-2. (2018)
- [18] Jun Xu, Xiangnan He, and Hang Li: Deep Learning for Matching in Search and Recommendation, <http://comp.nus.edu.sg/~xiangnan/sigir18-deep.pdf> (2018)
- [19] David Carmel, et al.: Product Question Answering Using Customer Generated Content - Research Challenges, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 1349-1350. (2018)
- [20] Manoj Kumar Chinnakotla and Puneet Agrawal: Lessons from Building a Large-scale Commercial IR-based Chatbot for an Emerging Market, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 1361-1362. (2018)
- [21] Sahin Cem Geyik, et al.: Talent Search and Recommendation Systems at LinkedIn: Practical Challenges and Lessons Learned, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 1353-1354. (2018)
- [22] Michael Skinner: Product Categorization with LSTMs and Balanced Pooling Views, https://sigir-ecom.github.io/ecom18DCPapers/ecom18DC_paper_9.pdf. (2018)
- [23] Rocio Canameres and Pablo Castells: Should I Follow the Crowd?: A Probabilistic Analysis of the Effectiveness of Popularity in Recommender Systems, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 415-424. (2018)
- [24] Jean Garcia-Gathright, et al.: Understanding and Evaluating User Satisfaction with Music Discovery, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp.55-64. (2018)
- [25] Hongyu Lu, Min Zhang, and Shaoping Ma: Between Clicks and Satisfaction: Study on Multi-Phase User Preferences and Satisfaction for Online News Reading, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp.435-444. (2018)