

会議報告

The 12th ACM Conference on Recommender Systems (RecSys 2018)

開催地: Parq Vancouver (バンクーバー, カナダ)

開催日程: 2018 年 10 月 2 (火) ~ 7 日 (日)

<http://recsys.acm.org/recsys18/>

1. RecSys 2018

ACM Recommender Systems (略称, RecSys) は, 推薦システムを専門とする国際会議である. 推薦システムとは, どの情報やものに価値があるかを特定するのを助ける道具である. この会議では, HCI (ヒューマンコンピュータインタラクション), 情報検索, および機械学習などの関連研究を, 推薦システムへの応用という観点からまとめている. Chris Anderson の “We are leaving the age of information and entering the age of recommendation.” の言葉に示されるように, 推薦システム研究は, 既存の分野から独立してきたとの考えに基づき, 推薦システム研究が盛んなミネアポリスにて 2007 年に第 1 回を開催した. これ以降, 2012 年の香港を除いて, 北米と欧州交互に開催されており, 第 12 回となる今回は北米で 6 回目の開催である. 筆者は 9 回目の参加で, 2010 年から続けて参加している. 開催地のバンクーバー市はカナダの西海岸最大の街であり, 国境を越えて南の US 側は IT 産業が集積したシアトル市がある.

例年より 1 日長く, 開催期間は 6 日であった. 初日はチュートリアル, 3 日間の本会議があり, 最後の 2 日はワークショップであった. 余談ではあるが, 1 日目の終了後に隣接のスタジアムでビヨンセのコンサートがあり混雑していた. オープニングで発表された参加登録者数は 42 개국から 816 人とこのことで, 新記録であった前回の参加者数 627 人から 3 割ほど大きく増えた. NIPS など 2018 年の機械学習の国際会議ではチケットの売切れが生じているが, この会議でもチュートリアルは売切れとなった. 国別参加者は US が最大で, 南米からの参加者が他の会議よりやや多く, ガボン, セネガル, ウガンダなどのアフリカ諸国からも参加者がいたようだ. インダストリとアカデミアの参加者は数年前まで同数程度だったが, 今回は 73% がインダストリからと大きな割合を占めた. 日本からの参加者は, 筆者の印象では前回とほぼ同じで 20 名前後で, アカデミアからは前回と同じく筆者を含めて 3 名程度であった.

スポンサーは 22 社と, 前回の 11 社から倍増である. ICML/KDD/NIPS など機械学習・データマイニング系の会議では Google/MS/Facebook の研究所がスポンサー

に並ぶが, ここでは Amazon だけであった. 推薦とつながりの深いコンテンツ配信・小売り企業, Spotify, Netflix, Pandora, Zalando などと並ぶ. 日本関連では, 前回に引き続きシルバーエッグが, 新たに NAVER/LINE もスポンサーになっていた.

2. 招待講演

招待講演は 3 件あった. 最初の講演は Google の Elizabeth F. Churchill によるもので, 社会の中での技術の運用についてであった. 五つの E と称して, Expedient (便利・実用), Exigent (需要がある), Explainable (説明できる), Equitable (公平), Ethical (倫理的) をシステムを設計する場合には考慮する必要がある. ここで, ちょっと「これらは Enigma (謎)」であるとオチを入れていた. 情報技術はよく使われるようになったが, いろいろ困る場合も生じていると. 中でも, Web 上で見られる偏向の問題を取り上げていた. 情報検索で著名な Ricardo Baeza-Yates の記事に基づき, これらの偏向は, 人間の認知のずれや, アルゴリズムの特性など, さまざまな原因で生じる. 加えて, 人間の推論は, インタフェースなど状況に非常に左右される. そこで, 人間の利用を前提として, 技術を解決に結び付けた technical computer-based システム (例: 文書を書くワードプロセッサ) ではなく, socio-technical システム (例: 本をつくるための出版システム) を考えることが重要であると述べた.

2 番目の講演は, ネットワーク分析で著名な Lise Getoor であった. グラフを対象とした, ラベル予測, リンク予測, 同一物判定を行う entity resolution の三つがあるが, これらはすべて論理式で表現できる. 確定的な論理では計算量や無矛盾性などの制約があるため, probabilistic softlogic (PSL) と呼ぶ確率論理を用いる. これは, 論理の変数を確率変数に, さらに述語にも重みを設定し, AND や OR を分布の最小値や最大値で表すものである. 理論上の下界や変分近似などを利用して効率的に解くことができると述べていた. この PSL を用いて, 推薦アルゴリズムを記述していた. 特に, 利用者がアイテムを評価したという単純な状況ではなく, 同時に利用者間のソーシャルネットワークでのつながりや, アイテム間の類似性に関する記述を加えるといった拡張が容易なことを協調していた. さらに, 持続可能性に配慮した商品に詳しい人の行動を利用してそうした商品を推薦したり, センシティブグループ間で出力を同じにする制約を加えて公平性を高めるさらなる拡張について述べた.

最後は, カナダの公共放送 Christopher Berry による

もので、カナダでの社会分断などに、ソーシャルネットワークなどが与える影響についてであった。

3. チュートリアル・ワークショップ

チュートリアルは5件で、今回は初日に当てられた。一つ目に聴講したのは、定性と定量を組み合わせた評価についてであった。正解率など、システム依存の定量的指標が主流であるが、ニーズの把握、仮説の精練、インタフェースやユーザ体験などはインタビューやアンケートによる定性的評価が必要になる。定性・定量評価をどのように組み合わせるかについて紹介した後、インタビューやアンケート調査にあたって、質問の仕方など注意すべき点をまとめていた。

二つ目に聴講したのは、系列推薦である。これは、直前にどの順番でアイテムを消費したかの情報に基づいて予測するものである。過去にどのアイテムを消費したかに基づく推薦は長期の嗜好を反映するが、この系列推薦は短期の嗜好を反映する。手法としては、頻出パターンマイニング、隠れマルコフや回帰結合ニューラルネットワークなどの系列モデル、そして時間方向を含むテンソル分解を用いる方法などがある。

11件のワークショップには最後の2日間が割り当てられた。深層学習に関するワークショップがあり数百人の参加者がいたようだ。その同時時間帯の開催であった公平性に関する **Responsible Recommendation** というワークショップの運営に筆者は携わっていた。これは、推薦や個人化技術の社会的影響や、その技術的対処について論じるワークショップである。ヨーロッパの公共放送における推薦技術の利用と、IEEEで議論されている倫理基準 **ethically aligned design** などに基づくシステム設計などについて招待講演を企画した。

4. インダストリアルセッション

このRecSysに特徴的なものとして、査読で採録する研究発表とは別に、企業の研究者や技術者が運用中のシステムに関する講演をするインダストリアルセッションがある。不動産検索の **Retailor** の講演は、不動産物件には一つとして同じ物件は存在しないという特性があるため、類似した物件を自己符号化器で近くに配置する特徴量抽出についてであった。ビデオ配信の **Hulu** は、三つ組形式の知識グラフを用いて説明可能な推薦を行う。このとき利用する知識が大まか過ぎたり、多過ぎたりすると良い説明にならないので、この点を工夫した。画像共有の **Pinterest** は、検索を個人化するために、利用者が保存している画像に関連した語を、検索語に加えてクエリ拡張を行う。ビデオ配信の **Netflix** はコンテンツの応じて映画のサムネイル画像を変更する。バンディットにより探索を行うが、利用者の混乱を避けるために頻繁には変更できない制約を考慮する。

5. 一般発表・受賞

発表種別はページ数で分かれたロングとショートがあり、どちらにも一部のものに口頭発表の機会が与えられる。ポスターは会議録に採録されないもので、意見交換などのための発表である。この会議では、単に予測誤差が下がればよいというのではなく、いかに利用者の情報要求を考慮しているかという点に着目している。

総投稿数は331件、ロングは181件のうち33件が採録、ショートペーパーは150件の投稿で38件が採録であった。採録率は、ロングは18.2%→20.8%→18%、ショートは20%→16.4%→25%とロングはやや厳しく、ショートはかなり緩くなった。日本からの本会議の発表はなかった。この会議は日本からの発表は弱く、過去にはロングは1本、ショート2本という状態である。

受賞についてまとめておく。

ロングベストペーパー **“Causal Embeddings for Recommendation”** : 推薦が利用者の行動に変化を与える因果関係を明示的に考慮した推薦モデル。

ショートベストペーパー **“Impact of Item Consumption on Assessment of Recommendations in User Studies”** : やはり推薦の影響に関するもので、利用者調査による検証を行った。

個人的に関心のあった一般発表をいくつかあげておく。今年は、新しい観点の発表が多数あり興味深かった。

- **Explore, Exploit, and Explain : Personalizing Explainable Recommendations with Bandits** : 既存の推薦は不適合なものは無視するが、それらのアイテムも利用者の嗜好を調査するのに利用する
- **HOP-Rec : High-Order Proximity for Implicit Recommendation** : 暗黙評価の場合での訓練事例の取得方法の改良。ベストショートの次点。
- **Calibrated Recommendations** : ジャンルごとのアイテムの比率などを反映した推薦。
- **Deep Inventory Time Translation to Improve Recommendations for Real-World Retail** : 商品の入替わりが頻繁な実店舗のために、今は扱っていない過去の類似商品の情報を利用。
- **How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility** : 人気バイアスにより推薦の多様性がなくなるかどうかをシミュレーションで検証。
- **Exploring Author Gender in Book Rating and Recommendation** : 書店の著者の性別分布と、推薦される本の著者の分布の差についての調査。
- **Unbiased Offline Recommender Evaluation for Missing-Not-At-Random Implicit Feedback** : オフライン評価で使うデータを一様ランダムに選ぶと、利用者が実際に選ぶ分布と違ってしまふので補正する。
- **Judging Similarity : A User-Centric Study of**

Related Item Recommendations : 利用者視点でのアイテムの類似性について利用者調査を行った。

6. おわりに

チュートリアル資料はホームページにて公開されており、また論文は今回からオープンアクセスになった。会議関連の Twitter の tweet は <https://togetter.com/li/1262109> にまとめておいたのでご参考にされたい。

2019 年はデンマークの首都コペンハーゲンで例年とほぼ同じ 9 月 16 ~ 20 日の開催と、2020 年は初めて南米ブラジルはリオ・デ・ジャネイロでの開催と告知があった。世界的には注目される会議にもかかわらず、日本からの発表や参加が、特にアカデミアから少なく寂しい現状だが、次回は発表や参加が増えればと思う。

[神嶌 敏弘 (産業技術総合研究所)]

The 8th Joint International Semantic Technology Conference (JIST 2018)

開催地: 淡路島夢舞台国際会議場 (兵庫県淡路市, 日本)

開催日程: 2018 年 11 月 26 日 (月) ~ 28 日 (水)

<http://jist2018.knowledge-graph.jp/>

1. 会議概要

Joint International Semantic Technology Conference (JIST) は、セマンティック Web, オントロジー, 推論, 知識グラフ, など, セマンティック技術を中心とした国際会議である。アジア地域におけるセマンティック Web の国際会議 Asian Semantic Web Conference (ASWC) の流れを汲んでおり, 2011 年に中国で開催された JIST 2011 より, 対象とするトピックをセマンティック技術一般に広げるとともに, 各国や地域で行われている会議と連携 (Joint) する形で毎年開催するようになった。日本での開催は, 2012 年の JIST 2012 (開催地: 奈良) に続き 2 回目で, 本学会のセマンティックウェブとオントロジー (SWO) 研究会との連携会議として開催された。

参加人数は約 95 名で 15 か国からの参加者があった。うち国外からの参加者は 33 名で, 韓国, 中国, タイなどのアジア諸国に加え, UK, ドイツ, チェコなどヨーロッパからの参加者も一定数見られた。

セッションの構成は, Technical Papers (口頭発表) に加え, 基調講演 1 件, 招待講演 1 件, Poster & Demo セッション, チュートリアル 2 件, ワークショップ 2 件が行われた。Technical Papers の論文投稿数は 75 件あり, うち 23 件が Regular Papers (採択率 31%), 6 件が Short Papers として採択された。

JIST 2018 から始まった新しい取組みとしては, Technical Papers セッションの一部を, Special Session として公募した点にある。提案者は Special Session ごと

に独自のプログラム委員会を組織し, そのセッションへの投稿論文の査読および採否判定を行える (最終の採否の判定は, JIST 2018 全体の Program Chair および Special Session Track Chair との協議のもとで決定する)。この仕組みにより, 最新の研究トピックに関する論文投稿の増加と, 新しい研究コミュニティの形成が促進された。

今回採択された Special Sessions は, Government Open Data および Semantic Web for Life Sciences の 2 件であった。

今回採択された Special Sessions は, Government Open Data および Semantic Web for Life Sciences の 2 件であった。

以下, それぞれのプログラムの概要について報告する。

2. 併設会議: 第 46 回セマンティックウェブとオントロジー研究会

JIST 2018 開催前日の 2018 年 12 月 25 日には, 併設会議として第 46 回セマンティックウェブとオントロジー (SWO) 研究会が開催された。今回は, SWO 研究会が主催する AI 技術に関するコンテスト「ナレッジグラフ推論チャレンジ」の発表会および授賞式も行われた。

ナレッジグラフ推論チャレンジは, 解釈可能性をもった AI 技術の開発を目指した技術コンテストである [川村 08]。第 1 回目である今回は「シャーロック・ホームズの短編小説『まだらのひも』をナレッジグラフ化したものを対象として, 事件の犯人を推論 (推理) し, 犯人と判断した理由 (動機や殺害方法など) の説明とともに示す」というタスクを競った。推論に用いる技術には制限を設けず, オントロジー, SPARQL クエリ, 述語論理, 記述論理, 機械学習など, さまざまな技術を利用した 8 件の作品が応募された。

これらの各応募作品についてのプレゼンテーションの後, 最終審査が行われ, 以下の受賞作品が決定した。

● 本部門

最優秀賞: 田村光太郎, 外園康智 (野村総合研究所)

優秀賞: 上小田中 411 (富士通研究所)

ベストリソース賞: FLL-ML (富士通研究所)

- アイデア部門

ベストアイデア賞：白松研 feat. 59 (名古屋工業大学)

これらの応募作品の詳細は、ナレッジグラフ推論チャレンジの公式サイト*¹に公開されている。

2. ワークショップ, チュートリアル, Poster & Demo

開催日の1日目には、2件のワークショップと2件のチュートリアル, Poster & Demoセッションが行われた。

1件目のワークショップは、

- The Fifth International Workshop on Practical Application of Ontology for Semantic Data Engineering (PAOS 2018)
- The 3rd International Workshop on Platforms and Applications for Social Problem Solving and Collective Reasoning (PASSCR 2018)

という二つのワークショップの合同開催として行われ、オントロジーやナレッジグラフの構築や利用、またそれに関わる技術について、7件の研究発表があった。

2件目のワークショップは、今回が初開催となる Workshop on Semantic Government Open Data (SGOD 2018) であった。このワークショップでは、同日の午後の Government Open Data をテーマとした Special Session で発表される論文の著者らによるパネルディスカッションが行われた。パネルディスカッションのテーマは

Challenges and Opportunities for Semantic Government Open Data

で、日本、タイ、台湾の3か国で行政のオープンデータに携わるパネルリストによる各国での課題の共有が行われた。

チュートリアルはワークショップと並行して以下の2件が行われた。

- How to Build Your Own Triplestore Using Table Data*²：山中遼太氏 (Oracle Corporation)
- Practical use of Knowledge Graph with Case Studies using Semantic Web Publishing tools*³：森川裕章氏, 小林賢司氏 (富士通研究所)

いずれのチュートリアルも、セマンティック技術の実利用に必要な基盤ソフトウェアの実践的な利用方法についてハンズオンを交えての解説がなされており、参加者の満足度も高かった。なお当日の資料は、それぞれ上述の脚注で示す GitHub レポジトリで公開されている。

同日の夜には、Poster & Demo セッションが行われ、一般応募より12件、スポンサー企業より1件の発表が



図1 Oscar 教授による基調講演の様子

あった。Reception を兼ねて飲食をしながらの発表形態を取ったこともあり、終了時間の間際まで熱心な議論が行われていた。

3. 基調講演および招待講演

2日目には基調講演として、マドリード工科大学 (Universidad Politécnica de Madrid) の Oscar Corcho 教授による Ontology Engineering at Scale for Open City Data Sharing という講演が行われた (図1)。本講演では、Oscar 教授がスペインで実施している都市に関するデータのオープンデータ化の取組みにおける経験を踏まえ、オントロジー工学がどのような役割を果たしているかを中心に紹介された。オープンデータ公開の重要性をさまざまな層の関係者に伝えるための工夫から、オントロジー構築方法論に基づいたツールによる技術的な支援に至るまで幅広い議論が行われた。また今回の JIST 2018 では、行政のオープンデータに関するワークショップや Special Session が実施されたこともあり、講演終了後も、アジア各国でオープンデータに取り組んでいるコミュニティとの交流が盛んに行われていた。

3日目の招待講演は、豊田工業大学の佐々木裕教授による Semantic Technology and Level 5 Autonomous Driving という演題で行われた。佐々木教授は同大学においてスマートビークル研究センターのセンター長を務めており、完全な自動運転を意味するレベル5の自動運転の実現のためにセマンティック技術が果たす役割について、同センターにおける取組みを交えて紹介された。現状の自動運転技術に残っている Open Questions とともに、オントロジーをはじめとするセマンティック技術によるアプローチの可能性が示され、講演後も多くの質疑応答が交わされた。

4. Technical Papers セッション

前述のように Technical Papers セッションは、通常の Research Track に加え、公募による Special Session Tack から構成される。Research Track は、

- Knowledge Graphs

*1 <http://challenge.knowledge-graph.jp/>

*2 <https://github.com/ryotayamanaka/jist2018-tutorial/wiki>

*3 <https://github.com/lod4all/jist2018-hands-on>

- Question Answering and NLP
- Data Management
- Ontology and Reasoning

の四つのセッションが行われ、オントロジーや推論、データ管理といった基盤技術に加え、ナレッジグラフや質問応答といった、セマンティック技術のより実践的な利用に関するものまで幅広い発表が見られた。

さらに、**Special Session Track** は、セマンティック技術の応用分野として多くの研究が行われている

- Open Government Data
- Semantic Web for Life Sciences

の 2 セッションが採択された。両セッションにおける論文投稿は、これまでの **JIST** に比べて大幅に増加しており、新しいコミュニティ形成を促進するという **Special Session** の役割が果たされたと思われる。

5. まとめ

今回で 2 度目の日本開催となる **JIST** であったが、論文投稿数、参加者数ともに、ここ数年来の最高記録を更新するという盛況のうちに終わった。これは、国内外におけるセマンティック **Web** 技術への関心の高まりの表れであると思われる。また、今回の **JIST 2018** の開催に当たっては、**SWO** 研究会幹事をはじめ、国内コミュニティの多くの方々に協力をいただいた。ここに、心より感謝の意を述べたい。

なお、次回の **JIST 2019** は、中国の杭州にて、**Chinese Knowledge Graph Conference** と連携しての開催が予定されている。

◇ 参 考 文 献 ◇

[川村 08] 川村隆浩, 江上周作, 長野伸一, 大向一輝, 森田武史, 山本泰智, 古崎晃司: 第 1 回ナレッジグラフ推論チャレンジ 2018 ~ 解釈可能な人工知能を目指して ~, 2018 年度人工知能学会全国大会 (第 32 回) 論文集, p. 1F101 (2018)

[古崎 晃司 (大阪大学)]