Report on the 11th International Workshop on Location and the Web (LocWeb 2021) and the 11th Temporal Web Analytics Workshop (TempWeb2021) at WWW2021

Dirk Ahlers
NTNU – Norwegian University of
Science and Technology
Trondheim, Norway
dirk.ahlers@ntnu.no

Erik Wilde
Axway
ewilde@axway.com

Marc Spaniol
University of Caen Normandy
Caen, France
marc.spaniol@unicaen.fr

Ricardo Baeza-Yates
Northeastern University
Silicon Valley, CA, USA
rbaeza@acm.org

Omar Alonso
Instacart
San Francisco, CA, USA
oralonso@gmail.com

Abstract

LocWeb and TempWeb 2021 were the eleventh events in their workshop series and took place co-located on 12th April 2021 in conjunction with The Web Conference WWW 2021. They were intended to be held in Ljubljana, Slovenia as a potentially hybrid event, but due to the pandemic, were fully moved online.

LocWeb and TempWeb were held as one colocated session with a merged programme and shared topics to explore similarities and introduce attendees to the two related and complementary areas. LocWeb 2021 explored the intersection of location-based analytics and Web architecture with a focus on on Web-scale services and location-aware information access. TempWeb 2021 discussed temporal analytics at a Web scale with experts from science and industry.

Date: 12 April, 2021.

Websites: https://dhere.de/locweb/locweb2021 and http://temporalweb.net/.

1 Introduction

LocWeb and TempWeb 2021 were the eleventh events in their workshop series and took place co-located on 12th April 2021 in conjunction with The Web Conference WWW 2021.[1] They were intended to be held in Ljubljana, Slovenia as a potentially hybrid event, but due to the pandemic, were fully moved online. This is therefore the second time that LocWeb and TempWeb did not have a physical location. Our experiences with that have already been discussed in the report for LocWeb 2020.

LocWeb and TempWeb were held as one colocated session with a merged programme to explore similarities and introduce attendees to the two related and complementary topics. Temporal and spatial analysis and retrieval share many similarities in how they are important enablers, but are often derived, second class citizens, that need additional efforts to play to their strengths. Combined spatio-temporal analytics is becoming more important as a method in Web mining. This was a good way to share insights, while keeping the individual workshops' characteristics. The similarities allowed us to easily build the shared programme, which we think worked well, under four sessions: User Behavior, Networks & Graphs, Analytics & Visualization, Time & Space.

LocWeb explored the intersection of location-based analytics and Web architecture with a focus on on Web-scale services and systems facilitating location-aware information access as well as on Spatial Social Computing on the Web. Workshop details, including most presentations, can be found on the workshop homepage[2] and on the workshop series homepage.[3] The previous LocWeb report for 2020 has been published in SIGIR Forum [Ahlers et al., 2020] as well as those for previous years.

The focus of TempWeb and the topics addressed are a "natural" match with The Web Conference. With digital content born almost two decades ago, the need for a more systematic exploitation of our digital cultural heritage as well as new analysis techniques, becomes evident. Hence, TempWeb is an ideal venue to exchange knowledge about temporal analytics at a Web scale with experts from science and industry. Further workshop details can be found on the workshop homepage.[4]

2 LocWeb Workshop Theme and Topics

LocWeb continues the main theme of Web-scale Location-Aware Information Access and spatial social computing. Subtopics include (i) geospatial semantics, systems, and standards; (ii) large-scale geospatial and geo-social ecosystems; (iii) mobility; (iv) location in the Web of Things; and (v) mining and searching geospatial data on the Web. The workshop encourages work describing Web-mediated or Web-scale approaches that build on reliable foundations, and that thoroughly understand and embrace the geospatial dimension through interdisciplinary perspectives.

The workshop's topics of interest were: Location-Aware Information Access, Spatial Social Behavior, Location-Aware Web-Scale Systems and Services, Geospatial Data Science for Social Good, Experience of virtual and physical place through online mediation, Physical and virtual/online

---

[1] https://www2021.thewebconf.org/
[2] https://dhere.de/locweb/locweb2021/
[3] https://dhere.de/locweb/
[4] http://temporalweb.net/

spaces for collaboration and information access, Urban Planning and Citizen Engagement, Geospatial aspects of Smart Cities, Location in the Internet/Web of Things, Open Geospatial Web Data, Location prediction in social media and the Web, Influence modeling and processing in static and dynamic spatio-social graphs, Evaluation of frameworks, metrics and algorithms, Large-scale Geospatial Ecosystems, Standards for Location and Mobility Data, Modeling Location and Location Interaction, Location-Based Social Networks, Geospatial Web Search and Mining, Mobile Search and Location-Based Recommendation.

3   TempWeb Workshop Objectives and Topics

The objective of TempWeb is to provide a venue for researchers of all domains (IE/IR, Web mining, etc.) where the temporal dimension opens an entirely new range of possibilities and challenges. The workshop's ambition is to keep shaping a community of interest on research issues resulting from the introduction of the time dimension in web mining and analysis. The maturity of the Web as well as the emergence of large-scale repositories of web content makes this very timely and as a result, a growing number of research projects and services are emerging that have this focus in common. Having a dedicated workshop will help, we believe, to take a rich and cross-domain approach to this continuous research with a strong focus on the temporal dimension.

To this end, TempWeb focuses on investigating infrastructures, scalable methods, and innovative software for aggregating, querying, and analyzing heterogeneous data at Web scale. Emphasis will be given to temporal data analysis along the time dimension for web data that has been collected over extended time periods. A major challenge in this regard is the sheer size of the data it exposes and the ability to make sense of it in a useful and meaningful manner for its users. It is worth noting that this trend of using big data to make inferences is not specific to web content analytics. A now-common strategy in post-genomic biology is to measure, quantitatively, the action of all (or as many as possible) of the genes at the level of the transcriptome, proteome, metabolome and phenotype, and to use computerized methods to infer gene function via various kinds of pattern recognition techniques. On the Web, to a large extent, we have also reached this point. Web scale data analytics therefore needs to develop infrastructures and extended analytical tools to make sense of these.

4   LocWeb Workshop Contributions

The workshop had one keynote and 3 full papers, forming an interdisciplinary combination of topics. Our PC members provided 3–4 reviews per paper on average, for a 60% acceptance rate. The proceedings [Ahlers et al., 2021] are available in the ACM Digital Library[5] as part of the overall WWW 2021 Companion. The papers are openly available when following the link from the conference proceedings page.[6] Slides are linked – where available – from the workshop homepage.

In her keynote, Flora Salim talked about *Learning Spatio-Temporal Behavioural Representations for Urban Activity Forecasting* [Salim, 2021]. She presented work around human activity

---

[5]https://dl.acm.org/doi/proceedings/10.1145/3442442#heading8
[6]https://www2021.thewebconf.org/wp-content/uploads/2021/07/toc_companion.html

patterns in cities for increased sustainability. She shared technical approaches around large-scale mutivariate sensor data from heterogeneous sources – including raster and point data, and with quality issues – contextual modeling, and spatio-temporal predictions and recommendations, for example trajectory prediction, visitor intent, and urban flow forecasting. She also talked about spatial data in these cases being based on timeseries data, semantic categorisation, and spation-temporal graph embeddings, as well as the question of granularity of both space and time for both data and predictions.

In *Cross-city Analysis of Location-based Sentiment in User-generated Text* [Stelzmüller et al., 2021], Christopher Stelzmüller, Sebastian Tanzer and Markus Schedl presented their approach to understand how citizens live and feel based on geolocated user-generated content and how that can be used in urban planning. Their work is novel in that it aims to understand relations between location and sentiment between cities around the world. That abstracts from individual sentiments and may give a "sentiment baseline" as a sentiment timeseries per POI category worldwide. They use POIs from the OpenStreetMap dataset and sentiment extracted from geolocated Tweets for the aggregated analysis. As in other work, the results depend on the quality of the available OSM data and the availability of Tweets about the respective POIs. They note that non-english speaking cities show weaker relations between POI and sentiment, and also that there is a bias making it more difficult to apply to developing countries, not only on spatial behaviour but also on assumptions on POI use and other cases that would be valuable to follow up.

In *Understanding & Predicting User Lifetime with Machine Learning in an Anonymous Location-Based Social Network* [Reelfs et al., 2021], Jens Helge Reelfs, Oliver Hohlfeld, Max Bergmann and Niklas Henckell use location analysis within the Jodel anonymous LBSN to predict user churn. They use a data set of 3 years from Saudi Arabia, where they examine the disjoint local communities They could show similarities across communities, with larger cities being more similar, and a countrywide model prodiving easier and suitable results.

Ovidiu Dan, Vaibhav Parikh, and Brian D. Davison look at improving IP address geolocation in *IP Geolocation using Traceroute Location Propagation and IP Range Location Interpolation* [Dan et al., 2021]. The authors are in a unique position as they can correlate 8.4 million IP addresses with measured locations from Bing mobile searches. This allows them to confirm that commercial IP geolocation databases are often inaccurate or lack sufficient granularity for certain use cases. It further allows them to have high granularity ground truth data available for their novel technique, which combines smart use of traceroute information with an interpolation of IP locations. They show how to propagate physical location information along traceroute paths and interpolate missing locations for certain IPs from the rest of their range, and also by the heuristic that differences in latency time related to geographic proximity. The work can be used with the CAIDA Archipelago (Ark) measurement infrastructure.

5 TempWeb Workshop Contributions

For its eleventh edition, TempWeb accepted four very positively reviewed submissions for oral presentation (acceptance rate of 50%). We interpret the high quality of the submissions and the frequent contributors to TempWeb, as indicators of an evolving community. It shows a clear sign of a positive dynamic in the study of time in the scope of the Web and evidence of the relevance of this effort.

The keynote talk was given by Prof. Ricardo Campos who is a Professor at Instituto Politécnico de Tomar in Portugal. Ricardo's talk on "Telling Stories through Timelines" gave a perspective on timelines as a central piece for organizing events and complex topics. Ricardo's talk covered summarization, construction of ground truth, evaluation and included demos that highlighted the main concepts.

The paper by Petrović, and Scholtes introduced "PaCo: Fast Counting of Causal Paths in Temporal Network Data" [Petrović and Scholtes, 2021]. In their work the authors address the problem, that network analysis cannot be directly applied to time-stamped network data. In particular, this problem arises time series data capturing, e.g., time-stamped social interactions, time-resolved cooccurrences or other types of relational time series. For example, the time-stamped links (A, B) and (B, C) the chronological ordering and timing determines whether a causal path from node A via B to C exists. To this end, they presented an efficient algorithm to count causal paths in time-stamped network data.

The second submission on "Learning to Persist: Exploring the Tradeoff Between Model Optimization and Experience Consistency" by Goldenberg et al. tackled issues in machine learning models and recommender systems in Web applications [Goldenberg et al., 2021]. To this end, they analyzed the tradeoff between providing the user with a consistent experience and suggesting an up-to-date optimal treatment. In order to overcome the challenges addressed, they proposed a hybrid approach utilizing a Persistence Key and Persistence Timeout. The proposed method provided a robust engineering solution and can be further calibrated for the best hyperparameters.

In "Analysis and Visualisation of Time Series Data on Networks with Pathpy" by Hackl et al. an open source software package was presented [Hackl et al., 2021]. The tool presented implements statistical techniques to learn optimal graphical models for the causal topology generated by paths in time-series data. In particular, pathpy supports network analysis and machine learning techniques for temporal data on networks. As it is built on robust, scalable, and easy-to-use data structures that are coherent with python's data science stack, pathpy is suitable for a variety of data science tasks in the context of temporal Web data.

Finally, the paper "Visualisation of Temporal Network Data via Time-Aware Static Representations with HOTVis" by Perriy and Scholtes presented a graph visualization tool for the visual analysis of temporal network data [Perri and Scholtes, 2021]. In particular, their software allows a time-aware static network visualizations that highlight the causal topology of temporal networks. Thus, vertices that can directly and indirectly influence each other are highlighted in order to support dynamic graph visualizations.

6   Conclusion and Future Directions

The discussion revolved around how time and space are often co-dependent aspects in much of the work presented here, and general experiences with the workshops. A recurring theme was data quality and location granularity, as well as robust system development.

The co-located format of the workshops was a valuable experiment to bring different topics and viewpoints together. Both workshops have been accepted for The Web Conference 2022.

Acknowledgements

References

Dirk Ahlers, Erik Wilde, Rossano Schifanella, and Jalal S. Alowibdi. Report on the Tenth International Workshop on Location and the Web (LocWeb 2020). *SIGIR Forum*, 54(1), June 2020. doi: https://doi.org/10.1145/3451964.3451972. URL http://sigir.org/wp-content/uploads/2020/06/p10.pdf.

Dirk Ahlers, Erik Wilde, Rossano Schifanella, and Jalal S. Alowibdi. LocWeb2021 Workshop – Proceedings of the Eleventh International Workshop on Location and the Web. In *Companion Proceedings of the Web Conference 2021*, WWW '21, New York, NY, USA, 2021. ACM. ISBN 9781450383134. doi: 10.1145/3442442.3451891.

Ovidiu Dan, Vaibhav Parikh, and Brian D. Davison. IP Geolocation Using Traceroute Location Propagation and IP Range Location Interpolation. In *Companion Proceedings of the Web Conference 2021 (LocWeb2021 workshop)*, WWW '21, page 332–338, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3451888. URL https://doi.org/10.1145/3442442.3451888.

Dmitri Goldenberg, Guy Tsype, Igor Spivak, Javier Albert, and Amir Tzur. Learning to persist: Exploring the tradeoff between model optimization and experience consistency. In *Companion Proceedings of the Web Conference 2021*, WWW '21, page 527–529, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3452051. URL https://doi.org/10.1145/3442442.3452051.

Jürgen Hackl, Ingo Scholtes, Luka V. Petrović, Vincenzo Perri, Luca Verginer, and Christoph Gote. Analysis and visualisation of time series data on networks with pathpy. In *Companion Proceedings of the Web Conference 2021*, WWW '21, page 530–532, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3452052. URL https://doi.org/10.1145/3442442.3452052.

Vincenzo Perri and Ingo Scholtes. Visualisation of temporal network data via time-aware static representations with hotvis. In *Companion Proceedings of the Web Conference 2021*, WWW '21, page 533–534, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3452053. URL https://doi.org/10.1145/3442442.3452053.

Luka V. Petrović and Ingo Scholtes. Paco: Fast counting of causal paths in temporal network data. In *Companion Proceedings of the Web Conference 2021*, WWW '21, page 521–526, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3452050. URL https://doi.org/10.1145/3442442.3452050.

Jens Helge Reelfs, Max Bergmann, Oliver Hohlfeld, and Niklas Henckell. Understanding & Predicting User Lifetime with Machine Learning in an Anonymous Location-Based Social Network. In *Companion Proceedings of the Web Conference 2021 (LocWeb2021 workshop)*, WWW '21, page 324–331, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3451887. URL https://doi.org/10.1145/3442442.3451887.

Flora D. Salim. Learning Spatio-Temporal Behavioural Representations for Urban Activity Forecasting. In *Companion Proceedings of the Web Conference 2021 (LocWeb2021 workshop)*, WWW '21, page 347–348, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3451892. URL https://doi.org/10.1145/3442442.3451892.

Christopher Stelzmüller, Sebastian Tanzer, and Markus Schedl. Cross-City Analysis of Location-Based Sentiment in User-Generated Text. In *Companion Proceedings of the Web Conference 2021 (LocWeb2021 workshop)*, WWW '21, page 339–346, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383134. doi: 10.1145/3442442.3451889. URL https://doi.org/10.1145/3442442.3451889.