Learning Service Semantics for Self-organization in Distributed Environments: Concepts and Research Directions

Graham Bent*, Geeth de Mel†, Raghu Ganti‡, Tom La Porta§,Gavin Pearson¶, Tien Pham‖
Sebastian Stein**, Leandros Tassiulas†† and Ian Taylor‡‡
*ETS, IBM Research, Hursley SO212JN, UK
†Hartree Centre, IBM Research, Warrington WA4 4AD, UK
‡IBM TJ Watson Research Center, Yorktown Heights, NY 10598, USA
§The Pennsylvania State University, University Park, PA 16802, USA
¶Defence Science and Technology Laboratory, UK
‖US Army Research Laboratory, Adelphi, MD 20783 USA
**University of Southampton, Southampton, SO17 1BJ, UK
††Yale University, New Haven, CT 06511, USA
‡‡Cardiff University, Cardiff CF24 3AA, UK

***Abstract*—A key challenge in performing analytics in distributed environments is to automatically compose services to dynamically match operational tasks to information requirements, accounting for impact, in a many-to-many temporally and spatially complicated and complex situations. In dynamic and agile environments, such as coalition environments, the state of the network and resources cannot be completely known in advance nor controlled due to the evolving nature of the network and constraints that may preclude partners from accessing complete state information about different parts of the system. In addition, there may be requests made to the system that have not been made before, requiring services to be created on the fly. Motivated by these observations, in this paper, we present a critical analysis of gaps in the state-of-the-art and our vision to address those through novel theoretical contributions. We envision that such formalized and theorized fundamentals will enable service elements to automatically configure themselves to perform analytic tasks based on user specified goals by taking account of context—be it system or user context.**

I. Introduction

A key challenge when seeking to efficiently and effectively perform analytics in resource-constrained distributed environments supporting temporally and spatially dynamic missions—such as occur in military and other first responder situations—is the ability to automatically compose complex services to dynamically match operational tasks to information resources, accounting for impact, in a many-to-many temporally and spatially complicated and complex situations. This requires the ability to understand the context of all the users of the information system, and take appropriate account when assigning—and re-assigning—resources to tasks. To provide such, a user context aware system requires the addressing of three critical research gaps:

1) deriving service compositions from the intent of generalist users;
2) conducting dynamic and distributed orchestration of the latent analytic capability across a resource-constrained, dynamic and distributed information infrastructure; and
3) understanding the criticality of the service to meeting the task's—hence the mission's—objectives and goals.

The resulting system then would be able to provide a birds-eye-view of the operating environment—especially in distributed coalition environments—by combining insights generated from multiple heterogeneous sources using coalition-wide shared services. Currently, much of this analysis is pre-planned with the data required to perform the analyses being collected from a variety of sensors and sources and processed in predefined stove-piped systems with the results being disseminated to the users.

To provide greater flexibility, micro-service architectures in which new applications are created from available component services [1], [2] are being more widely adopted. While micro-services provide greater flexibility, micro-service composition techniques face significant challenges when they need to be combined into complex services [3], [4]. The problem is exacerbated in the coalition context due to a variety of reasons: (1) services are distributed and typically poorly annotated, thus making meaningful compositions impractical—if not impossible; (2) requirements—i.e., user intentions—vary depending on the context, thus interpretation is hard; (3) requirements evolve over space and time, thus latent analytic capabilities keep propping up; and (4) assessment of criticality—and impact—of services to requirements is made difficult as the environment is dynamic—i.e., services and requirements are migrating through space and time, thus having varying utility in a given snapshot of time.

Motivated by the above observations, in this paper, we present our vision to realise a system which can adapt to the

Authors are organised in alphabetical order and have made equal contributions. †Geeth de Mel is the corresponding author and please direct any questions at geeth.demel@uk.ibm.com.

Dstl/CP107670. Content includes material subject to © Crown copyright (2018), Dstl. This information is licensed under the terms of the Open Government Licence except where otherwise stated. To view this licence, visit http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3 or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk

varying requirements in the coalition operating environment such that it can autonomously compose services, especially from micro-services, while taking into account the impact and the utility of the composition—and the resulting assignments.

The rest of the paper is structured as follows: in Section II we motivate our work through an illustrative scenario. In Section III we provide a high-level architecture of our system and discussed its desired features in Section IV. We then conclude the paper in Section V by discussing the work done to date to realise this vision and by sketching current research directions.

II. Illustrative Scenario

Let us assume that a mission is taking place in an urban environment to monitor the movement of a high value target: part of the mission is a task to localized the target. The target is in a specific vehicle and monitoring the movement involves closed-circuit television (CCTV) systems, infrared sensors, acoustic identifies, location services, and social media posts. Here we assume a system wherein specific models for detection, identification, and distinction can be uploaded to devices so that information processing services could generate insights that are relevant and meaningful for the mission in progress. We now assume that the high value target has reached an underground car park, thus resulting in having to monitor vehicles as well as humans so that the target is not missed. Figure 1 depicts the object detection with specialized image classification models.

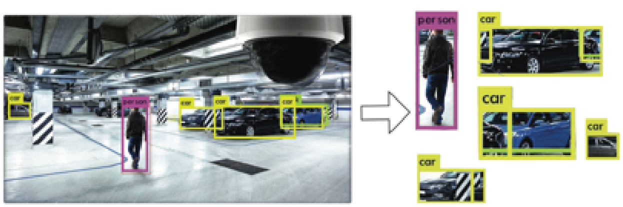

Fig. 1. Object detection with specialized image classification models.

As soon as the target vehicle moves into the underground car park, video and acoustic analysis get created by combining with location and predictive services for real-time surveillance. This situation demonstrate some specific requirements for a system as the one we mentioned in Section I–i.e., (1) the analytic has to be created and performed in real-time as the system with the car park was never configured to perform this task in the first place; (2) supports user intent matching to services so that it will autonomously solve the problem of which models and processing services are needed to best extract information requested; (3) impact and the utility of service creation needs to be considered as it involves intensive use of compute and communication resources; and (4) multiple data services (e.g., videos, trained models) and processing services (detection, localization, tracking) need to come together in a seamless manner to achieve the objectives of the task, thus the goal of the mission.

In the next section, we present the high-level architecture of this envisioned system, especially with respect to a brain-inspired model.

III. High-level Architecture

Our goal for the system is to determine the best analytic services to meet requests, and then to allocate those services to the outstanding requests; for services, we consider both data and logic to perform analytic functions associated with distributed coalition settings. Specifically, we use a brain-inspired model—i.e., a model where local data and processing services are co-located—to model services found in coalition environments. The intuition here is that the human brain seamlessly process localised data to create local insights and combine in a way that makes global inference possible [5]: similarly, if we consider services in a coalition environment as parts of a brain that generates local insights, combining them in a meaningful manner may provide insights about the evolving situation in the coalition environment. Figure 2 depicts our grand vision for the system.

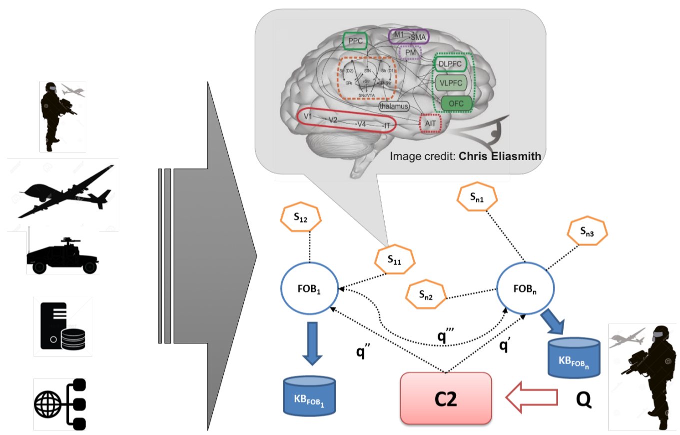

Fig. 2. Self organising services in dynamic distributed environments.

We consider a coalition operating environment which consists of many information providing sources—e.g., soldiers with connected gear, manned and unmanned vehicles, mobiles and wearable devices, and internet connected sources which may be relevant to the task at hand. We envision that these sources observe, gather, and create knowledge about their environment by means of processing services associated with them and expose the local knowledge by means of micro- and complex- services; S_{ij} in Figure 2 depicts such services. These services then come together to create more global insights through services such as FoB_i and even more insights through services such as $C2$.

Once the services are up and running, we provide novel querying capabilities through the system as shown in Figure 2. Given a query, we determine the closest match to analytic processing and data available in the system; the result will be a set of possible solutions, and an achieved utility based on the request and the closeness of match with the solutions. We then consider the potential strategies employed by the service requesters as they may report their utility strategically, and we provide assignment mechanisms that are robust to such strategies.

We also consider that requests may be satisfied by more than one resource in some cases, and that in other cases there may be no perfect match between possible services and requests, thus yielding to the need for best partial matches with respect to mission context, especially in mission-critical situations. We also consider that resources are constrained physically (e.g., energy, bandwidth) and logically (e.g., user intentions, policy prohibition, and so forth) having varying utility in dynamic coalition context, thus making it impractical—if not impossible—to satisfy every request. To address such we aim to deploy general matchmaking and resource allocation algorithms with respect to the perceived utility of services. Last but not least, the system explicitly considers a coalition environment in which some services may be shared, and partners may not fully or honestly disclose the utility of their missions.

In the next section, we discuss some desired features of this architecture in detail.

IV. Desired Features for Self-organisation

In order to create a system in which services are self-organising, the system needs to provide means to service modules such that they are self-describing, and that can self-discover and self-assemble to meet the required goals while adhering to the constraints of the coalition. Such an approach would enable the system to discover data-to-services and services-to-services relationships so that effective service negotiation is made possible in disconnected, dynamic environments. To achieve this goal of self-description, we propose to investigate techniques to automatically learn concepts/properties of different domains, create and extend representations with respect to new knowledge, and have mechanisms to dynamically expand knowledge bases with new information.

A. Self-describing Analytics

Existing techniques in self-description, discovery, and assembly require service providers and consumers to agree on a set of definitions that describe the service and how to consume them—i.e., application programmers must manually compose these descriptions and actively publish them for consumers [6], [7]. We propose to examine ways to automate this process by boosting existing descriptions through observational characterization. In order to characterize services we will look at feature-extraction techniques that extract characteristic behaviours from passively collected traces of communication channels between service consumers and services. For example, we can collect time-series data from communication channels between cooperating entities, feed these into a neural network, extract the feature-vector from the hidden layers, and feed these to a clustering algorithm [8]. Furthermore, to characterize local data—especially at the edge—one can explore more efficient techniques for latent process modelling and change-point detection in time-series data. The former can generate topic labels that are used to summarize data sets so that matching can take place efficiently, and the latter can approximate a probabilistic mixture model that can be used to compare service behaviours.

We also note that in a coalition setting, individual service meta-data, descriptions, logging will not follow pre-set syntax nor semantics. Thus, we need to describe individual services in a format that will allow us to learn from historical usage data, and dynamically compose new services. The format should allow computing distances between services, that reflect their semantic similarity. We propose to represent service by a vector which can be learned over time to capture its semantics. For instance, word2vec [9] is an example that allows learning semantic vectors to represent words : node2vec [10] is a scheme that captures the locality and neighbourhood information of a node within a graph. This is analogous to our system requirements—i.e., learn vector representation of services to build complex services via vector compositions. To do so, we not only need to capture the functionality of the service, but also the (1) workflow graph (may have loops, self-loops); (2) composability of services; (3) security, policy restrictions from coalition partners; and (4) cost of invoking the service, thus requiring a scheme that goes much beyond known vector learning algorithms.

Once the vector representations for services are learnt, creating composition will come down to learning sequences of operations, which can be done effectively by developing algorithms based on Recurrent Neural Networks (RNNs) [11]. However, typical RNN problems do not consider branching issues but in complex service compositions, there might be several branches that may or may not merge in the future. Determining the most efficient methods for handling branching and merging, and how to automatically learn semantic representations of such complex services an emerging challenge to be addressed for a system like ours.

B. Semantically-aware Analytic Matchmaking

In prior work, there is extensive work on matching sensors of various types to competing missions with different requirements [12]. Specifically, we used semantic reasoning to match assets to requests [13], developed utility functions to capture the fit of sensor assignments to missions [14], and dealt with budgets [15], [16]. There are two main differences with the proposed work here when compared to the previously mentioned work. Firstly, the aforementioned work focuses on sensors and missions in a sensor network and thus relies on physical devices and locations, whereas in this work we are looking at a high level of abstraction that deals with information distilled from information resources processing in a network. Thus, we are not assigning a single, or multiple, distinct physical devices to detect known phenomena based on their location, but are instead (1) determining if processing services and data exist that can be used together to answer a query; and (2) determining the closeness of the information that we distil to the original request. There are two entities that need to be matched (processing services and data), and the compatibility of the data and processing service must also be matched.

Secondly, the semantic matchmaking work in previous research only considered complete matches—though partial matches were considered w.r.t. individual resources, they were bundled together as resource packages to fully cater for the needs of requests—for resource requests w.r.t. to a predefined ontology. However, in coalitions analytic domain, we envision that partial matches will have to be considered, especially w.r.t. the utility provided by them, even when composite services are put together to satisfy requests. This is because parties involved in a coalition may compute the completeness—or value of the information—provided by analytic very differently; also, the parties involved in the coalition will naturally temper the analytic and the information coming from them w.r.t. the perceived provenance and trust matrices.

Therefore, in this system, we aim to use such properties to develop contextual ranking mechanisms for analytic, especially when there are multiple matches—full or partial—to a request. Additionally, we perceive an ontology—or the domain model—associated with the matchmaking work to be an agile representation—i.e., we will consider means in which the ontology is expanded based on the selections made by the analytic users through techniques such as reinforcement learning and active learning mechanisms. The first has already been applied in centralized service repository research but we aim to investigate techniques on doing so when the analytic are deployed in distributed coalition environments where systems do not possess a birds-eye-view of the analytic space. Active learning techniques will enable our mechanisms, e.g., to have a conversation with human experts, especially when there are unknown situations, thus enhancing the future experiences of the matchmaking algorithms. Thus, these techniques will enable us to introduce the instincts human experts bring into the resource and request matching to algorithmic space too.

C. User-intent Driven Intuitive Service Composition

Command staff, at all levels of command[1], will receive orders from higher command, conduct an estimate of the situation[2], create a plan[3] and issue orders (which will usually include a graphical representation of the scheme of manoeuvre and a synchronisation matrix (about timing of activities)). Thus, it can be envisaged that the activities associated with creating the plan, orders, scheme of manoeuvre etc could result in the creation of *user-intent driven* service compositions. These might include such things as (1) Activity in *Area of Interest* detection, classification and alerting service; (2) Enemy course of action prediction service; and (3) Readiness and availability of communications/logistics given Blue course of action.

These would, in turn, need to be composed of more basic services such as those within the *Warfighters Information Services* framework [17] such as (1) Object tracking service; and (2) Own force location service. Each of which would in turn spawn a need for communication, data and analytics services; these would be composed from the available micro-services[4].

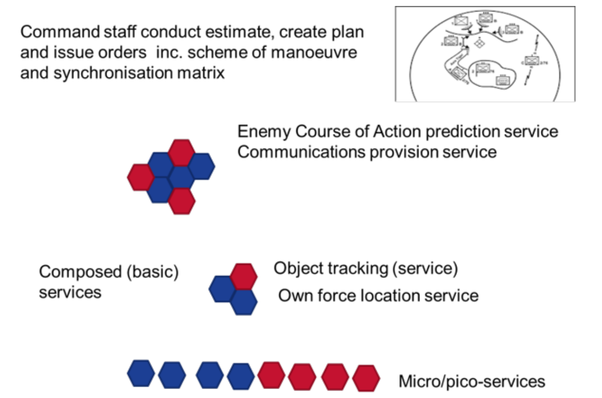

Fig. 3. Accounting for Impact on the Mission: Understanding the criticality of the service to meeting the mission/task goals/objectives

As noted above, another key challenge is how to manage the priority of the different service requests; and how to make this be based upon ‘user-intent’ and the criticality of the service to meeting the mission/task goals/objectives.

Ethnographic study *Understanding Commanders Information Needs* [18] has identified that commanders use full staff briefings to ensure the staff understands which items of information are of potential significance and time sensitive (i.e. which need to be brought urgently to the commanders and wider staff attention, and which can wait till the next scheduled brief). Empirical studies by Moffat *et. al* [19] and Medhurst *et. al* [20], [21] have shown that is it possible to use decision-games to develop a probabilistic model of the value of different categories of information to the decisions made.

D. Strategy-aware Resource Allocation

An overbearing assumption in state-of-the-art resource allocation algorithms is that they assume service requesters truthfully report their requirements and mission utilities, even if this means their analytic tasks are delayed or rejected. This is unrealistic in a coalition environment, where requesters are drawn from across many different coalition partners; although they share the same broad objectives, requesters also pursue the interests of their own internal group. Thus, they should be treated as self-interested agents that may behave strategically (e.g., by misreporting analytics priorities or constraints); additionally, different coalition partners may view requests as having different utility, thus adjusting their responses.

To address this key shortcoming, we plan to use techniques from the field of mechanism design [22], which provides tools

[1]This breaks down at below Company level; as the company is the lowest level of a military organisation with a dedicated headquarters (which is only 2-5 people).

[2]The ‘seven questions’ in current UK doctrine; the exact method may vary by country.

[3]The commander will review the plan.

[4]If they could not be provisioned then the system would automatically provide feedback up to the ability to provide the ’user-intent driven’ services.

for incentivising agents to truthfully report their utilities. While there is existing work that has applied mechanism design to resource scheduling [23], [24], it does not deal with the challenges inherent in coalition environments, including the high dynamism of analytics tasks, rapidly changing network compositions and demand, and balancing coalition objectives with the limited view of self-interested task requesters. Specifically, the system will enable following capabilities through the new science of mechanism design in coalition settings: (1) We will develop online mechanisms for resource allocation in coalitions—i.e., we will develop new mechanisms that do not know in advance what tasks will be submitted and need to make scheduling decisions immediately as new tasks arrive; (2) We will investigate adaptive learning mechanisms for resource allocation in coalitions—i.e., we will consider realistic coalition operations where accurate statistical knowledge of demands (as assumed by existing work10,) is not known in advance and where conditions change rapidly over time. Thus, there is a need for allocation mechanisms to learn from and adapt to the prevailing situation; and (3) We will also balance global mission objectives with local interests—i.e., mechanisms typically aim to maximise the social welfare within a system. However, this may not always be appropriate as for example the coalition may sometimes wish to prioritise the tasks of particular agents (e.g., if those are involved in a critical or a very time-sensitive mission).

Additionally, some requesters may have a poorly calibrated view of their own requirements, which could lead to their tasks being allocated incorrectly. To address this, our system will look at global control mechanisms that can calibrate and adjust task priorities between competing agents.

E. Real-time Service Allocation

One of the challenges in service allocation, especially in the coalition context, is that they cannot be solved in an offline manner since the demand for resources is not known a priori. As a result, most service assignment algorithms are heuristics-based and cannot maximize using the current-state-of-the system. A notable exception, however, is the case of the max-weight scheduling algorithm [25] (with variants for data analytics [26]), which can maximize the use of resources in a system without statistical knowledge of the analytics arrival processes. However, optimality comes at the price of an exponential increase of the delay. This issue is critical in many systems, but in particular when delay sensitive applications compete for resources with throughput-intensive applications that are also delay tolerant. As a result, system utilization is usually sacrificed (i.e., over provisioning) in order to obtain a good delay performance.

The importance of providing services with low delay is widely acknowledged by both industry and academy, and has received a lot of attention in recent years [27], [28], [29], [30]. For instance, in [15], authors show that low delay variability is crucial for providing many of their analytics—including web search and augmented reality—and identifies causes of the delay in their systems. These include multiple flows competing for computation and communication resources , and the use of multiple queueing layers interconnecting resources. The approaches in [16,17] take the delay issue to the extreme, and propose heuristic algorithms that work without queues, thus enhancing the performance in terms of delay; however, in such work, the system utilization attributes are not clear w.r.t. resources.

In view of all this, we believe it is essential to establish the mathematical foundations that allow us to maximize the use of resource while providing low delay to real-time analytics. The technical approach we will follow consists of combining stochastic optimization techniques from control theory (max-weight [1]) with deterministic interior-point [9] algorithms in convex optimization. These two approaches are in marked contrast to each other: while max-weight algorithms make decisions in an online manner and create congestion, interior-point methods produce no congestion but they can only be implemented in an offline manner. Hence, they cannot be used in real system. Our goal is to bring the knowledge of max-weight algorithms to interior-point methods, and design a new class of algorithms—which we call approximate interior-point methods—that can handle stochasticity, are resilient to system perturbations, and are able to provide low-delay.

F. Dynamic and distributed orchestration

Last but not least, am import feature of the system is to be able to allow the dynamic and distributed orchestration of the latent analytic capability available across the infrastructure. The lack of any centralised awareness of the state of the information system means that individual service components need to be able to decide to join a service pipeline and that such distributed self-forming service pipelines must not become co-mingled by mistake.

V. CONCLUSION

In this paper, we have provided a critical analysis of the challenges when performing analytics in distributed environments, especially on automatically composing complex services to dynamically match operational tasks to information resources, accounting for impact, in a many-to-many temporally and spatially complicated and complex situations. A potential system to address these challenges are proposed and potential techniques are identified to address those challenges. Currently, we conducting our research in multiple strands to achieve this common vision. Specifically, we are investigating on (1) mechanisms to derive service compositions from the intent of generalist users; (2) strategies to conduct dynamic and distributed orchestration of the latent analytic capability across a resource-constrained, dynamic and distributed information infrastructure; and (3) means to understand the criticality of the service to meeting the task's—hence the mission's—objectives and goals. In the near-to-mid future, we plan to publish our systematic finds, simulated and real-world evaluation results specific to the components of the system in reputed venues.

Acknowledgement

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

[1] C.-C. Huang, Q. Chen, Z. Wang, R. Power, J. Ortiz, J. Li, and Z. Xiao, "Spartan: A distributed array framework with smart tiling." in *USENIX Annual Technical Conference*, 2015, pp. 1–15.

[2] A. Ludwig, S. Schmid, and A. Feldmann, "Specificity vs. flexibility: On the embedding cost of a virtual network," in *Teletraffic Congress (ITC), 2013 25th International*. IEEE, 2013, pp. 1–9.

[3] C. Esposito, A. Castiglione, and K.-K. R. Choo, "Challenges in delivering software in the cloud as microservices," *IEEE Cloud Computing*, vol. 3, no. 5, pp. 10–14, 2016.

[4] N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, and L. Safina, "Microservices: yesterday, today, and tomorrow," in *Present and Ulterior Software Engineering*. Springer, 2017, pp. 195–216.

[5] C. Eliasmith, *How to build a brain: A neural architecture for biological cognition*. Oxford University Press, 2013.

[6] E. Wittern, J. Laredo, M. Vukovic, V. Muthusamy, and A. Slominski, "A graph-based data model for api ecosystem insights," in *Web Services (ICWS), 2014 IEEE International Conference on*. IEEE, 2014, pp. 41–48.

[7] M. Á. Rodríguez-García, R. Valencia-García, F. García-Sánchez, and J. J. Samper-Zapater, "Ontology-based annotation and retrieval of services in the cloud," *Knowledge-Based Systems*, vol. 56, pp. 15–25, 2014.

[8] W.-H. Lee, J. Ortiz, B. Ko, and R. Lee, "Time series segmentation through automatic feature learning," *arXiv preprint arXiv:1801.05394*, 2018.

[9] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[10] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2016, pp. 855–864.

[11] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[12] T. L. Porta, C. Petrioli, C. Phillips, and D. Spenza, "Sensor mission assignment in rechargeable wireless sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 10, no. 4, p. 60, 2014.

[13] M. Gomez, A. Preece, M. P. Johnson, G. De Mel, W. Vasconcelos, C. Gibson, A. Bar-Noy, K. Borowiecki, T. La Porta, D. Pizzocaro, H. Rowaihy, G. Pearson, and T. Pham, "An ontology-centric approach to sensor-mission assignment," in *International Conference on Knowledge Engineering and Knowledge Management*. Springer, 2008, pp. 347–363.

[14] H. Rowaihy, M. P. Johnson, D. Pizzocaro, A. Bar-Noy, L. Kaplan, T. La Porta, and A. Preece, "Detection and localization sensor assignment with exact and fuzzy locations," in *International Conference on Distributed Computing in Sensor Systems*. Springer, 2009, pp. 28–43.

[15] M. P. Johnson, H. Rowaihy, D. Pizzocaro, A. Bar-Noy, S. Chalmers, T. La Porta, and A. Preece, "Frugal sensor assignment," in *International Conference on Distributed Computing in Sensor Systems*. Springer, 2008, pp. 219–236.

[16] ——, "Sensor-mission assignment in constrained environments," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 11, pp. 1692–1705, 2010.

[17] S. Bray, "A framework for warfighter information services–using the concept of a virtual knowledge base," in *Proceedings of the 15th International Command and Control Research and Technology Symposium (ICCRTS)*, 2010.

[18] J. P. Kahan, D. R. Worley, and C. Stasz, "Understanding commanders' information needs," RAND CORP SANTA MONICA CA, Tech. Rep., 1989.

[19] J. Moffat and J. Medhurst, "Modelling of human decision-making in simulation models of conflict using experimental gaming," *European Journal of Operational Research*, vol. 196, no. 3, pp. 1147–1157, 2009.

[20] J. Medhurst, I. Stanton, H. Bird, and A. Berry, "The value of information to decision makers: an experimental approach using card-based decision gaming," *Journal of the Operational Research Society*, vol. 60, no. 6, pp. 747–757, 2009.

[21] J. Medhurst, I. Stanton, and A. Berry, "Risk taking by decision makersusing card-based decision gaming to develop models of behaviour," *Journal of the Operational Research Society*, vol. 61, no. 11, pp. 1561–1571, 2010.

[22] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge university press, 2007.

[23] S. Stein, E. H. Gerding, A. Rogers, K. Larson, and N. R. Jennings, "Algorithms and mechanisms for procuring services with uncertain durations using redundancy," *Artificial Intelligence*, vol. 175, no. 14-15, pp. 2021–2060, 2011.

[24] V. Conitzer and A. Vidali, "Mechanism design for scheduling with uncertain execution time." in *AAAI*, 2014, pp. 623–629.

[25] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE transactions on automatic control*, vol. 37, no. 12, pp. 1936–1948, 1992.

[26] A. Destounis, G. S. Paschos, and I. Koutsopoulos, "Streaming big data meets backpressure in distributed network computation," in *INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, IEEE*. IEEE, 2016, pp. 1–9.

[27] G. J. Chen, J. L. Wiener, S. Iyer, A. Jaiswal, R. Lei, N. Simha, W. Wang, K. Wilfong, T. Williamson, and S. Yilmaz, "Realtime data processing at facebook," in *Proceedings of the 2016 International Conference on Management of Data*. ACM, 2016, pp. 1087–1098.

[28] M. P. Grosvenor, M. Schwarzkopf, I. Gog, R. N. Watson, A. W. Moore, S. Hand, and J. Crowcroft, "Queues don't matter when you can jump them!" in *NSDI*, 2015, pp. 1–14.

[29] J. Perry, A. Ousterhout, H. Balakrishnan, D. Shah, and H. Fugal, "Fastpass: A centralized zero-queue datacenter network," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 307–318, 2015.

[30] S. Kulkarni, N. Bhagat, M. Fu, V. Kedigehalli, C. Kellogg, S. Mittal, J. M. Patel, K. Ramasamy, and S. Taneja, "Twitter heron: Stream processing at scale," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. ACM, 2015, pp. 239–250.