Will Distributed Computing Revolutionize Peace? The Emergence of Battlefield IoT

Tarek Abdelzaher*, Nora Ayanian§, Tamer Basar*, Suhas Diggavi‡, Jana Diesner*, Deepak Ganesan†,
Ramesh Govindan§, Susmit Jha‖, Tancrede Lepoint‖, Ben Marlin†, Klara Nahrstedt*, David Nicol*,
Raj Rajkumar**, Stephen Russell††, Sanjit Seshia¶, Fei Sha§, Prashant Shenoy†, Mani Srivastava‡,
Gaurav Sukhatme§, Ananthram Swami††, Paulo Tabuada‡, Don Towsley†, Nitin Vaidya*, and Venu Veeravalli*
*University of Illinois at Urbana-Champaign, Urbana, IL 61801
†University of Massachusetts, Amherst, MA 01003, ‡University of California, Los Angeles, CA 90095
§University of Southern California, Los Angeles, CA 90089, ¶University of California, Berkeley, CA, 94720
‖SRI International, New York, NY 10165, **Carnegie Mellon University, Pittsburgh, PA 15213
††U.S. Army Research Laboratory, Adelphi, MD 20783

***Abstract*—An upcoming frontier for distributed computing might literally save lives in future military operations. In civilian scenarios, significant efficiencies were gained from interconnecting devices into networked services and applications that automate much of everyday life from smart homes to intelligent transportation. The ecosystem of such applications and services is collectively called the *Internet of Things (IoT)*. Can similar benefits be gained in a military context by developing an IoT for the battlefield? This paper describes unique challenges in such a context as well as potential risks, mitigation strategies, and benefits.**

I. Introduction

Can distributed computing reduce the human and economic toll of military operations? A recent US study estimates the total cost of military operations, since the beginning of the millennium, at 4.79 *trillion*,[1] or about $15,000 per person in the US, making it one of the largest expenses for the nation. While much of that cost, as the title suggests, is attributed to armed conflicts, in this paper we draw examples from the entire gamut of military operations, including such missions as non-combatant evacuations, peacekeeping operations, and humanitarian missions (e.g., disaster response).

From a computational perspective, reducing the *human* cost of military operations, expressed in such factors as injuries and loss of life, can be attained by a form of *automation* that reduces collateral damage (e.g., via smarter sensing, control and actuation) and decreases the need for physical presence of individuals in risky environments.

Similarly, reducing the *economic* cost can be attained, among other means, by improved early response. For example, an earlier and better-informed response to a humanitarian need (such as the recent post-hurricane crisis in Puerto Rico) would generally lead to a lower long-term operation cost. This improved response is often referred to as *readiness*.

Since most world population lives in cities (according to recent studies that show urban populations tipping the balance since 2008), we further envision military operations that are increasingly carried out in urban contexts. This environment further exacerbates the challenges of reducing human and economic cost.

The question to the research community that we focus on in this paper therefore becomes: what research advances can simultaneously meet the aforementioned two intertwined requirements; namely, (i) increase the level of automation and (ii) improve response (or readiness) in military operation scenarios, especially ones carried out in urban spaces? At a high level, the two requirements may appear to be in conflict. Delivering a faster and more appropriate response to unexpected conditions calls for improved human decision-making, whereas automation aims to remove humans from the loop. How can one advance both simultaneously? This paper explains how distributed computing solutions can be a significant part of the answer, thereby contributing to one of the most impactful research directions in computing in the next decade, both in terms of humanitarian and economic benefits.

In order to simultaneously meet the requirements of increased automation and readiness, this paper describes future networks of *battlefield things* that can recognize, empower, and properly carry out *commander's intent* in a safe, responsive, and resilient manner. These networks execute battlefield services that aim to meet mission requirements in unpredictable, fast-changing, distributed adversarial environments. An investigation into the analytical foundations of the Internet of Battlefield Things (IoBT) has recently started under a new US Army research program by the same name.[2] Below, we offer a perspective on the underlying challenges, risks, mitigation solutions, and prospective impact.

The simultaneous need for increasing automation and improving response has significant implications on computing challenges. Intuitively, a faster response requires a *shorter decision loop*. Consider, for example, a non-combatant evacuation operation, where civilians must be safely removed from a zone of increased or impending hostility. The situation is highly dynamic. New information updates arrive in real-

[1] https://www.usnews.com/news/articles/2016-09-09/report-post-sept-11-wars-have-cost-479-trillion

[2] https://www.eurekalert.org/pub_releases/2017-10/uarl-bit100417.php

time regarding the situation in affected locales. This arriving information may impact decisions such as evacuation routes and deployment plans for protection forces. One needs to optimize the ability of decision-makers to arrive at faster better-informed decisions.

In military contexts, decisions are hierarchical. The hierarchical nature of decisions reduces the speed of response as authorizations to carry out actions must arrive through an appropriate chain of command. As a result, actions are delayed and, by the time they are carried out, might already be based on stale information. A recent departure from the strict hierarchical structure in military thinking was the adoption of a *command by intent* doctrine; a paradigm that empowers subordinate units to exercise more initiative and autonomy in the context of a mission, without waiting for explicit directives from upper echelons. In this paradigm, a commander specifies their intent (such as evacuating non-combatants along safe routes), leaving it largely to the subordinate units to fill-in the details. The paradigm shortens the decision loop, enabling better local exploitation of fleeting opportunities, as well as improving decisions by acting faster (and, hence, on more up-to-date data).

The original command by intent doctrine largely envisions humans as the mission execution agents. When combined with the need for increasing automation, arising from the desire to reduce human cost, exercising autonomy and initiative becomes more challenging. How can teams of distributed machines and humans execute command by intent? Indeed, how does one manage increasingly autonomous smart assets in ways that allow them to improvise and exhibit initiative to meet mission needs, while offering some assurances on aggregate behavior, and while reacting and re-configuring around disruptions and failures at different scales caused by a harsh environment or a determined adversary? We note that the above is a core computer science problem.

In the rest of this paper, we elaborate what we mean by IoBT and break down the above problem into more specific research directions in distributed computing.

II. IoBT Challenges

A recent US Army vision[3] defines IoBT as a set of interdependent and interconnected entities that can include sensors, actuators, devices (computers, weapons, vehicles, robots, human-wearables, etc), infrastructure (networks, storage, processing elements), algorithms (on-node, in-network), information sources, and humans. These entities (i) are dynamically composed to meet multiple missions, tasks, or goals; (ii) operate autonomously and autonomically; and (iii) perform intelligent battlefield services (such as capture/process data, predict behaviors/activities, and effectuate the physical environment) in order to enable predictive analytics and deliver intelligent command and control.

The application domain has challenging characteristics that collectively distinguish battlefield IoT from civilian scenarios. These characteristics include:

- *Diverse missions, tasks, and goals*. An IoBT might be specifically created and adapted to meet a mission, complete a task, or accomplish a goal. There will likely be many networks operating simultaneously, possibly competing for resources. The possible tasks are diverse. They include wide area persistent surveillance, tracking a dispersed group of humans and vehicles moving through cluttered environments, monitoring cities for protests, disaster relief operations, or monitoring physiological and psychological state of soldiers, to name a few. Tasks are not expected to start or end simultaneously, and new tasks may emerge as others are being executed.
- *Highly dynamic, mobile, and resource-constrained environment*. Many networks will be forward-deployed and will consist of disadvantaged assets with limitations on energy, power, storage, and bandwidth. Fixed infrastructure may not be available, posing limits on computing, and communications. Services delivered on the IoBT must operate under these severe constraints and will often need to support tasks with limited time availability.
- *Extreme heterogeneity*. The variety of things available to an IoBT is immense, ranging from very capable devices and simple disposable ones. In addition to the various classes of things described above, a network will contain a mixture of entities that include military devices controlled by the military (which we henceforth call *blue* assets), adversarial devices (which we call *red*), and devices controlled by neutral entities (which we call *gray*). Co-existence and co-deployment of commercial Internet of Things (IoT) devices and networks with purposefully built, certified, and carefully controlled military devices and networks will be required. Networked entities will thus have a wide range of security levels and capabilities that must be accommodated.
- *Varying scale*. IoBTs will be deployed in a wide variety of places and domains, usually in contested environments. One extreme is the highly dense and cluttered mega-city environment. Another extreme is sparse terrain with limited entities and gaps in sensor coverage and networks.
- *Contested and adversarial environments*. Many IoBTs will be deployed with limited physical security and will include entities in the IoBT that are owned and controlled by the adversary. IoBT must be protected from a variety of sophisticated and persistent threats. Security measures must be taken to protect against determined intelligent adversaries. Analytics must deal with conflicting and deceptive data, and identify adversarial activity.

Within the above context, we divide the envisioned high-level IoBT functions into three key types: (i) synthesis (of desired capabilities), (ii) adaptive/resilient execution, and (iii) learning. We further define these capabilities as follows:

- *Assured synthesis:* In response to a need, how can one

[3] https://www.arl.army.mil/www/pages/3050/IOBT-Program-Announcement-AmendmentII.pdf

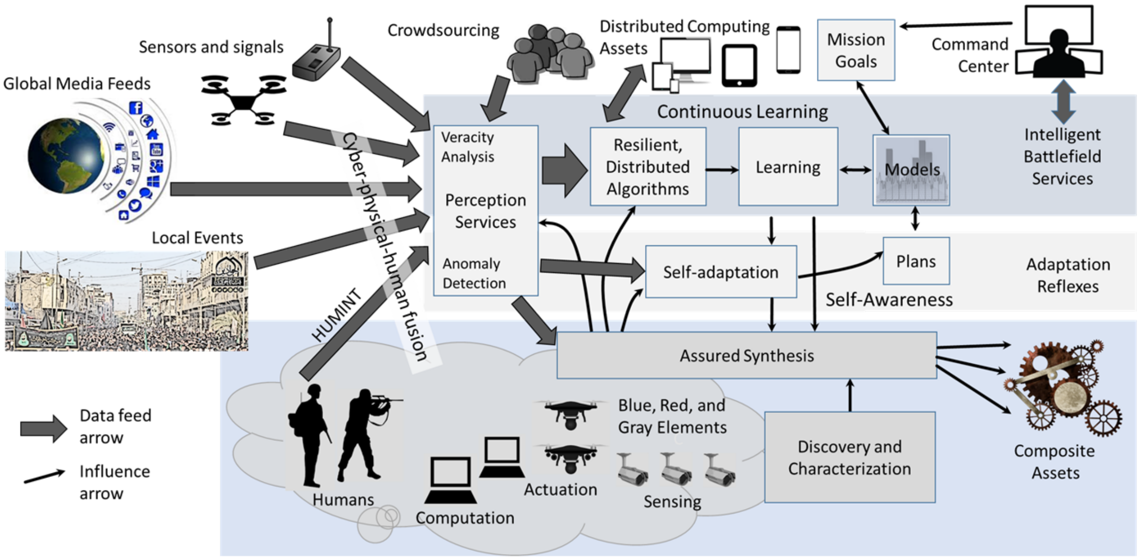

Fig. 1. An Internet of Battlefield Things

quickly synthesize a capability that meets the need? What assurances can be made about the behavior of the synthesized capability? Such an assured synthesis capability contributes to readiness by improving the speed and efficacy of response to unexpected conditions.

- *Adaptive/resilient execution:* Once synthesized, how can one quickly adapt and reconfigure capabilities in response to a disruption, threat or loss? How does one adapt to the environment in order to optimize success metrics? Akin to instinctual reflexes, this fast adaptation capability is needed to handle sudden disturbances, setbacks and opportunities, while executing a mission.
- *Learning:* How does one manage and support intelligent and learning battlefield services that accumulate knowledge, learn from experience, and continually improve outcomes against determined adversaries?

Furthermore, security is a crosscutting concern. One must ensure that all of the above is safe and secure, even in the presence of malfunction, infiltration, and partial loss of resources. The interactions among the above capabilities in an IoBT context are depicted in Figure 1.

III. Challenge 1: Assured Synthesis

Future missions will exploit IoBTs made of thousands or tens of thousands of blue/military, red/adversary, and gray/citizen nodes, with a wide range of capabilities (Figure 2): from tiny occupancy sensors to drones with three-dimensional Radar and LiDar sensors; from small on-board compute devices to powerful edge clouds with GPUs; and from actuators capable of modifying the environment in some way, to humans with powerful (albeit biased) perception, cognition, and action capabilities. Furthermore, future missions will need to be exceedingly *agile*: mission goals and needs may not be known until just before mission execution, and mission planners may not be able to (without the aid of automated tools) recruit and construct, at short timescales, IoBTs with sufficient resources to satisfy mission needs. The large scale of IoBTs implies continuous churn, so discovery and composition solutions will need to be robust to failure or removal of assets as a normal operating regime.

A research challenge is therefore to develop methods and fundamental limits underlying the recruitment and composition of IoBT resources, including potentially adversarial ones, into composite assets with sufficient sensing, compute, and communication capacities to satisfy mission needs and constraints. Recruitment, composition and reconfiguration of such assets must meet two fundamental needs. First, it should be possible to assemble (or re-assemble, for example, upon damage) composite assets comprising an IoBT of possibly 1,000s to 10,000s of nodes on demand and within an appropriately short time (e.g., minutes, if needed), despite high component heterogeneity, large scale, and presence of adversaries. Second, the aggregate properties of the composite, including timeliness, performance/functionality, security, and dependability, must be formally assured in an appropriately quantifiable and operationally relevant manner, subject to well-understood assumptions. Several important scientific advances can contribute to the solution space. For example:

- Algorithms for discovery of gray/red nodes using side channel emanations.
- Algorithms and mathematical foundations for rapid top-down synthesis of mission-specific IoBT functions, offering composable assurances of correctness and composable assessments of risk.
- Algorithms and theory for exploitation of physical dynamics of sensor observations to enable secure and resilient state-estimation and control in the face of data contamination.
- An understanding of fundamental information-transfer and capability limits of composite IoBTs.

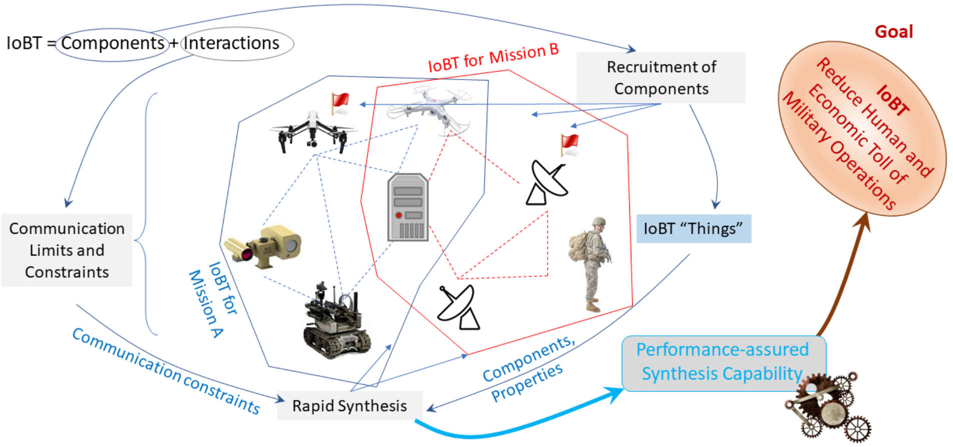
IoBT = Components + Interactions
IoBT for Mission B
Recruitment of Components
Goal
IoBT
Reduce Human and Economic Toll of Military Operations
Communication Limits and Constraints
IoBT "Things"
IoBT for Mission A
Communication constraints
Performance-assured Synthesis Capability
Components, Properties
Rapid Synthesis

Fig. 2. Synthesis of Large-scale Composite IoBTs

Below we detail some of the above ideas. The goal is to enable not only efficient composition of assets that meet mission needs, but also improved quantification of risk, brought about by a better understanding of assurances guaranteed by such composite assets. Hence, *disciplined* initiative may be exercised in mission execution (as opposed to poorly-informed "gambling").

A. Recruitment

One must develop methods to discover assets (sensors, actuators, and humans), characterize their capabilities to meet mission goals (and/or their potential threats, in case of gray/red nodes), and recruit the appropriate assets into an IoBT. Capabilities of assets include properties such as performance, reliability, trust and security. We envision two qualitatively different research threads in this context: discovery and characterization of cyberphysical assets, and discovery and characterization of human assets. A critical challenge underlying both of these threads is the resilience of discovery and characterization to adversarial behavior.

Cyberphysical assets: While discovery and characterization of fixed wired assets has made impressive strides, algorithms and limits of the discoverability and characterizability of cyberphysical, mobile, and wireless assets at large scale remains less well-studied. Cyberphysical assets can potentially be discovered through cyber-discovery techniques (probing, snooping, fingerprinting based on unique traffic characteristics), but this alone is insufficient, since these devices have qualitatively different characteristics from wired devices: they may be intermittently connected, so may not consistently respond to probes or emit traffic; they have several connectivity options (cellular, Wifi, Bluetooth), so may not appear at consistent topological locations and may not be amenable to fingerprinting based on traffic characteristics; they contain a variety of sensors whose characterization is crucial for IoBT missions; and they may move frequently, so their discovery needs to be continuous. The above difficulties make cyberphysical asset discovery and characterization an interesting research problem.

Human assets: Recent research created theory-driven frameworks for human-in-the-loop sensing, we call social sensing [1], [2]. Social sensing offers estimation-theoretic and system identification-based approaches to characterize human sources. Properties of interest include human reliability and bias. Theoretical models and effective and scalable algorithms were developed that automatically discover ground-truth from possibly noisy, biased, linguistically ambiguous, and conflicting claims provided by various information sources [3], [4]. Fact-finding algorithms were developed to characterize reliability of sources on social media (e.g., bloggers on Twitter) and compute confidence in results [2]. These approaches need to be extended to offer a foundation for identifying and characterizing human components that work in various capacities within an IoBT. Humans (for example, members of the local population in a city) can collectively offer sensing, actuation, or control in an IoBT context making it important to properly model and account for their behavior. A related important issue is the discovery of resources (e.g., sensing, computational, and actuation resources) available to such humans that can be leveraged for asset composition.

Resilience to adversarial behavior: Adversarial nodes (either humans or cyberphysical assets) may contaminate data, disrupt discovery and characterization, or perform impersonation attacks. New methodologies must be developed for robust asset discovery and characterization in highly dynamic and unknown environments.

B. Composition

Given a high-level description of a mission's goal, one must automatically synthesize a composite asset comprising (a) an IoBT network, and (b) associated distributed software services that can, together, satisfy the missions needs.

This capability requires addressing at least three challenges: (i) *automatic reasoning from goals to means* to derive requirements and constraints from high-level goal specifications, (ii) *network composition* that satisfies the requirements and constraints, and (iii) *functional composition* for generating distributed services and controllers that achieve the mission goals in a scalable manner.

Reasoning from goals to means: First, given a high-level goal such as "track a collection of insurgents and report on their activities and rendezvous points within a certain geographic area", one must address such questions as: what sensors and actuators are needed to achieve the goals, what in-network compute elements must be present to achieve the desired latency, and what network capacity and resilience must exist to fulfil mission information extraction needs? Efficient analytically-founded approaches must be developed to reason about *functional* requirements. The ability for such on-the-spot fast composition or improvisation of highly-customized capabilities to mission goals brings about the next degree of agility and adaptation to military operations in contested environments. Prior work on macroprogramming [5]–[7] and network service composition [8], [9] offers instances of deriving functional behavior from high-level specifications. Top-down IoBT synthesis requires generating the network topology, dealing with adversarial elements, and reasoning about computation capabilities for information processing and machine learning.

Scalability: Given a set of network capacity, computation, latency and resilience requirements, one must *synthesize* a near-optimal network meeting those requirements from discovered IoBT assets. Formalisms must be developed to address this problem, such as constraint satisfaction [10], and optimization-theoretic approaches [11]. Essentially, these approaches search discovered IoBT nodes to determine subsets that optimally satisfy the requirements (under some optimization objective such as using fewest discovered nodes, or fewest adversarial assets, etc). However, the potential search space is very large because of the heterogeneity of sensors, actuators and compute elements. Thus, clever solutions must be developed to address tractability. They may include a judicious choice of constraints to reduce search space, or perhaps a hiearchical problem decomposition that exploits independence relations between subproblems.

A key challenge in composition lies in understanding fundamental constraints brought about by *time* and *bandwidth* limitations on information transfer among the composed components, as well as optimizing information transfer subject to these constraints.

IV. Challenge 2: Adaptive Reflexes for IoBTs

The next challenge is to develop theoretical foundations, models, and methods for autonomy, autonomicity, and self-awareness of composite IoBT assets, such as those synthesized as described above. In biological systems, *reflex theory* states that complex behavior can be attained (and thus explained) through the combined action of individual reflexes that have been chained together. Can one develop a parallel theory for IoBTs that offers foundations, models, and methods for autonomous, autonomic, and (more generally) self-aware behavior in the face of distribution, scale, dynamics, heterogeneity, resource constraints, and presence of adversaries? The following key scientific advances are needed:

- A unifying theory of self-aware adaptation inspired by multidisciplinary foundations borrowing from technical areas such as self-stabilizing algorithms, information theory, and adaptive control.
- Game theoretic foundations for hierarchical decomposition of global goals into objectives for distributed subordinate subsystems that jointly achieve the overall goal.
- Quantifiable assessment metrics for self-aware and self-adaptive systems.

In the long term, the area develops the foundations for exercising disciplined *initiative* while carrying out mission orders. It does so by decomposing high-level mission goals into specific objectives for subsystems, while allowing for local adaptation at the subsystem level that nevertheless ensures quantifiable compliance, in aggregate, with mission goals. The resulting system architecture is depicted in Figure 3.

A. Foundations for Self-aware Adaptation

An exciting and intellectually stimulating undertaking is to develop a *unified theory of adaptation* in self-aware systems. These foundations can gain inspiration from understanding adaptive behaviors traditionally studied *across multiple disciplines*, and putting them under the same analytic foundation. Importantly, the unified treatment should allow investigation of *aggregate behavior* when individual, *largely heterogeneous*, adaptive components *interact* (whose adaptation algorithms, today, are studied in *different* non-interacting disciplines). This would be the case, at scale, in the heterogeneous and dynamic environment of IoBTs. Of particular interest is the situation where some of the adaptive components are malicious, aiming to derail the aggregate converged behavior of the composite. Prior work has shown that uncoordinated interactions of adaptive components, *even when aimed at meeting the same goal*, can result in unexpected consequences and severe performance loss [12]. This challenge of mitigating undesirable collective behaviors of heterogeneous adaptive systems was recently emphasized in the outcomes of a Dagstuhl Seminar on Self-aware Computing as one of the key challenges for the field [13]. Indeed, although elements of self-* properties pervade the most diverse areas of engineering (and special instances are studied in a stove-piped fashion in different disciplines), there is currently no unifying theory of self-aware adaptation applicable to the heterogeneous environments of IoBTs. An important challenge would be to build such a theory, together with appropriate assessment metrics, as one key foundation of *self-awareness*, thereby empowering distributed components to adapt in coordinated ways that respond to local stimuli (such as failures) while meeting global requirements on behavior of the collective.

Multi-disciplinary foundations: Individual examples of adaptation in distributed and centralized systems are plentiful. For example, *self-stabilizing algorithms* adapt to maintain an invariant by triggering corrective action, when the invariant is

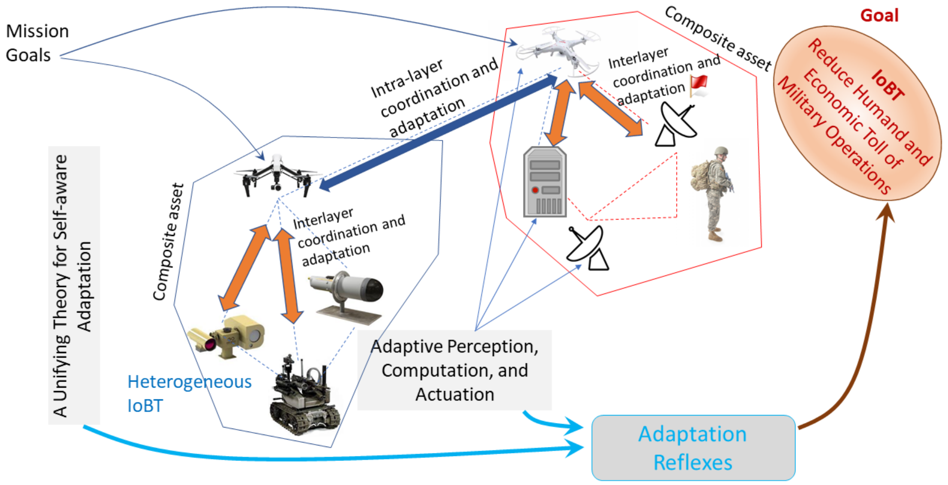
Mission Goals
Intra-layer coordination and adaptation
Composite asset
Interlayer coordination and adaptation
Composite asset
Interlayer coordination and adaptation
Goal
IoBT
Reduce Human and Economic Toll of Military Operations
A Unifying Theory for Self-aware Adaptation
Heterogeneous IoBT
Adaptive Perception, Computation, and Actuation
Adaptation Reflexes

Fig. 3. Adaptive, Self-aware IoBTs.

violated, to cause the system to satisfy the invariant again. Similarly, in *error correction* codes, a notion of goal can be understood as (avoidance of) constraints imposed in the codewords (e.g., all the codewords must have an even number of ones). If these constraints are violated, a decoder is able to correctly decode the message by re-enforcing the constraints imposed by the code. We can thus see error correction as a form of adaptation upon the occurrence of a mismatch between the received message state and the goal. As the last example, consider *adaptive control*, where the notion of goal is expressed in a desired state of the physical system being controlled. During execution, the sensed data and actuation commands are constantly checked for consistency of model-based predictions, as well as against the control goal. In case of mismatch, the model is adapted to better agree with the data, and new control commands are issued accordingly, to better achieve the goal. Here, adaptation lies in revising both the plant model and control command.

While the previous examples are from different scientific disciplines (distributed computation, information theory, and control, respectively), use different mathematical tools, and offer different self-* properties, they all implicitly share the notion of *self* that encapsulates state, models, actions, and goals, and that adapts its actions and models as needed, such that its goals are met. Can this simple principle serve as the cornerstone of a new unifying theory of self-aware adaptation, addressing the challenges brought by IoBTs, including massive heterogeneity and scale, operation in contested environments, and seamless interaction between humans and machines? How does one incorporate security as a native part of this theory? Armed with a capability for self-aware adaptation, one may enable construction of IoBTs that use mission orders and exercise initiative to adapt locally while ensuring invariants in line with commander's intent.

Operationalizing agent interactions: Another important question is: how to design distributed coordination and control mechanisms governing the interactions between agents starting with the desired objectives and ending with the specific interaction mechanisms? A possible answer might leverage game-theoretic advances in the design of *multi-level dynamic games* that offer provable convergence guarantees on the end result. Namely, by suitably choosing agent objective functions, one may be able to guarantee that the interactions between the multiple agents in the battlefield will converge to an equilibrium in which the desired objectives are met. The necessary distributed coordination and control between agents do not need to be explicitly designed, but rather naturally result from each agent seeking to optimize its given objective function. This approach is an analytic embodiment of command by intent, carried out by machines. Namely, objectives (i.e., objective functions) are given to subordinates, while granting them the freedom to exercise initiative on how best to meet these objectives. The analytic framework for designing such objectives would allow devices in an IoBT to follow the same doctrine. The process of optimizing the local objective function of an agent can be seen as the operational counterpart of self-aware adaptation, since the objective function encodes the notion of self and how it relates to the agent's objectives, actions, and models. The approach is scalable because each agent is empowered to perform the operations needed to optimize its objective function *without explicit coordination with other agents*, thereby minimizing overhead.

B. Adaptive Perception, Computation, and Actuation

To offer concrete degrees of freedom expressed in "adaptation knobs", one should also develop a suite of adaptation algorithms for distributed systems in response to mission requests, for example, by reconfiguring (i.e., moving) or rebinding to different sets of physical resources and computational algorithms in response to changing context.

Adapting perception resources: Key to adapting perception resources is the existence of some form of redundancy (i.e., ability to perform functions in several alternative ways). While discovering redundancies in processing capabilities of network nodes is relatively easy, the same is not true for sensing and actuation capabilities, which depend not only on the type of sensors and actuators but also on their precise coupling to the physical world. Discovering these redundancies requires learning (and adapting, as necessary, via passive observation and active probing) a model of how a distributed and dynamic set of sensors and actuators couple to each other through the physical world. For example, seismic sensing may be used when smoke or other phenomena render visual tracking unreliable, or when connection is lost with the camera due to a wireless jamming attack.

Adapting computing and communication resources: In IoBT environments, there will be significant dynamics. Groups of devices will be composed dynamically, and may need edge or backend processing or storage. Critical data from IoBT device groups will also be highly dynamic – depending on where interesting events are being seen, there will be a need to dynamically (re)allocate computing and network resources to meet time and capacity constraints [14]. Adversaries in IoBT environments could potentially inject data or noise to saturate processing resources, starve communication, or isolate information sources. Resource allocation algorithms will be needed that can (i) dynamically reallocate heterogeneous resources at the edge, network core, and backend to handle rapidly changing situations in connectivity and needs, (ii) scale resource allocations to match workloads that exhibit high spatial and temporal variability, and (iii) prevent any subset of IoBT devices (including attackers) from saturating cloud processing and communication resources.

Adapting actuation and control resources: An important aspect of heterogeneity is diversity in how tasks are accomplished. In fact, diversity is well documented as a way to improve the performance of human workgroups. Studies have shown repeatedly that diverse groups outperform homogeneous groups [15]–[18]. Thus, instead brittle controllers designed with fixed assumptions, one may design novel controllers that are parameterized differently but adapt their parameterization by observing their neighbors, so that the system self-adjusts to the environment.

V. Challenge 3: Learning and Intelligent Battlefield Services

Complementing the discovery and composition of heterogeneous goal-driven IoBTs and endowing them with reflex-like adaptation functionality, this challenge is concerned with supporting the learning capabilities of IoBTs. Such support functions may include reliable information gathering in adversarial settings, large-scale distributed processing, and analytic services for fast and efficient machine intelligence in distributed and adversarial settings, as shown in Figure 4. Numerous fundamental challenges must be addressed within this context to overcome existing gaps and limitations, as follows:

- Theories and algorithms to ensure trustworthiness of data gathered in adversarial environments, prior to its input to fusion engines, learning systems, and decision support tools.
- Theories and algorithms for distributed learning and analytics over heterogeneous data in the presence of network adversities and adversarial compromise.
- Theories and algorithms for improving learning safety, robustness, and cost.

These challenges are further elaborated below.

A. State Assessment and Diagnostics

A useful service would be to develop mathematically well-founded algorithms for assessment of state and quality of data inputs to fusion engines, learning systems, and decision-support tools, especially in the presence of possible lack of trust in data sources or sensors. This leads to multiple research problems of a diagnostic nature, as discussed below.

System diagnostics: A key challenge in the complex environments of IoBTs is to diagnose distributed *system health*. This problem is especially daunting given the possible lack of observability of many system components. Health, therefore, needs to be inferred (and damage, if any, assessed) *without direct component observation*. In communication networks, this problem is sometimes known as network tomography [19]; discovery of latent network structure (or structural compromise) from a sample of end-to-end observations [20]–[22]. IoBTs feature significantly increased heterogeneity, compared to traditional networks, as well as a faster pace and scale of dynamics (compared to the more stable Internet topology). The higher prevalence of adversarial elements in contested environments also adds to the challenge. New diagnostic algorithms are needed that support heterogeneous IoBT graphs, address scalability, distribution, streaming, and high rate of topological dynamics, while operating in a harsh adversarial environment.

Information diagnostics: A different problem is the identification of bad (human or physical) sources, erroneous sensor signals (due either to malicious activities or to faults), as well as automatically explaining causes of anomalous behavior [23]. This work may leverage prior advances in data fusion, truth-finding, and reputation algorithms, among other possible frameworks. The topic offers interesting challenges. On the one hand, in fast-paced situations that involve many simultaneously moving parts, attention is a bottleneck. It should be directed to situations that deserve it the most. For example, it should be directed more to *anomalies* as opposed to *normal conditions*. On the other hand, in the presence of failures and noisy data, anomalous inputs might be the result of noise or misinformation. Therefore, focusing on them would be a distraction. How does one develop services that properly direct attention to situations/information that demand it the most, even in the presence of noise, failures, bad data, malicious adversarial inputs, and other possibly intentionally-designed distractions?

B. Learning Services

A key challenge is to develop a theory and algorithms for *learning* that are more suitable for the distributed, adversarial IoBT environments. Several questions arise in this context. For

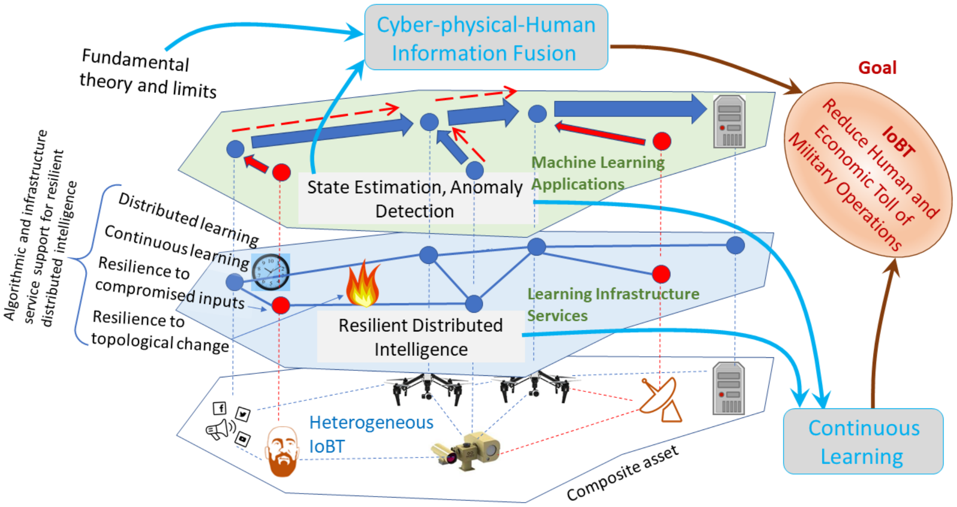
Cyber-physical-Human Information Fusion
Fundamental theory and limits
Goal
IoBT Reduce Human and Economic Toll of Military Operations
State Estimation, Anomaly Detection
Machine Learning Applications
Algorithmic and infrastructure service support for resilient distributed intelligence
Distributed learning
Continuous learning
Resilience to compromised inputs
Resilience to topological change
Learning Infrastructure Services
Resilient Distributed Intelligence
Heterogeneous IoBT
Composite asset
Continuous Learning

Fig. 4. Intelligent Battlefield Services

example: (i) how to *automatically distribute* machine learning computations over heterogeneous networks of sensing and compute nodes subject to loss of network connectivity; (ii) how to make distributed learning *continuous and never-ending* to handle the needs of ever-changing battlefield environments, and (iii) how to make distributed learning *resilient* to adversarial compromise and network adversities? We elaborate below.

Distributed learning: The current state of the art in distributed machine learning-based data analytics assumes that models and algorithms are run over secure, reliable networks and that all required features are available all the time. Another significant limitation of current distributed machine learning systems is that they are only marginally tolerant of heterogeneous hardware configurations. Data-parallel implementations of machine learning algorithms perform the same computation over all compute nodes by definition [24], [25]. This precludes the use of truly heterogeneous networks of nodes from wearables to compute clusters whose storage, memory, and compute capabilities may differ by many orders of magnitude. In the context of IoBTs, new theories and algorithms are needed that accommodate heterogeneity, and tolerate a wide array of failures and adversarial compromises of learning nodes. Indeed, a study is needed into the resilience of distributed learning to *adversarial* change. For example, what is the impact of time-varying topology (such as that caused by failures due to an adversary) on the correctness and convergence of distributed learning algorithms? Resilient learning algorithms need to be developed for intelligent battlefield services executed on IoBTs.

Continuous and robust learning: In systems that learn blindly without proper contextualization, new information can often erase previously learned knowledge [26]. Informally speaking, “appropriate behavior” must be contextualized. One must learn what is proper for each context. Importantly, the system must learn the different relevant underlying contexts automatically. *Adverarial* attacks may supply malicious inputs (i.e., inputs modified to yield erroneous model outputs) [27]. In an IoBT environment, an adversary may control red/gray nodes and observe (hence, label) our digital and physical reactions to inputs of its choice. The continuous scale at which learning happens in an IoBT environment makes it susceptible to fine-grained modifications of inputs. Formal methods are required to verify the correctness of systems that include machine learning components in the presence of adversarial training data.

Optimizing cost of learning: In a dynamic IoBT, especially where communication is short and significant network dynamics and heterogeneity are present, the cost of network learning could be non-negligible. One would need to explicitly account for the cost of deploying and using the network. In recent work, information theoretic results were developed for networks where the cost of learning is explicitly accounted for, demonstrating that one might activate different network topologies based on the trade-off between network learning and communication [28]–[33]. This work may inform design of dynamic IoBTs that self-configure to jointly optimize both learning cost and decision making accuracy.

Ensuring learning safety: A long-term goal is to develop formal verification technology for supporting development of high-assurance adaptive and self-aware learning systems. This is difficult due to the very large set of reachable states in learning systems, easily going beyond the limits of current verification technology, as well as the presence of feedback loops that break compositionality properties, making reasoning about safety of the overall behavior very challenging. Since traditional approaches are inadequate for dealing with the complexities of learning systems, novel methodologies are needed that might rely on runtime monitoring, certificate-based verification, and combining data-driven and symbolic techniques. A particularly promising approach here is based on simulation-driven verification and runtime monitoring of suitable properties of the system [34]. Preliminary results also exist on extending symbolic reasoning engines that have had significant impact on techniques in Formal Methods, to establish safety bounds on data-driven learned models [35].

VI. Discussion

The challenges discussed above set the stage for a new era of battlefield IoT services that simultaneously increase automation and improve quality of response in unexpected military scenarios, thereby reducing both the human and economic toll of military operations. A key cross-cutting theme that underlies many of the discussed challenges is the tussle between dependability, autonomy, and learning.

For example, the work requires reconciliation of the ability to offer assurances (e.g., via formal methods) with the ability to autonomously learn [36]. One must be able to verify composable, intelligent, adaptive, and learning systems. A significant research problem is to develop abstractions and algorithms to model, verify, and synthesize systems that incorporate machine learning (such as deep neural networks) in safety-critical and mission-critical perception and decision-making tasks. Recent work [37] describes challenges for applying formal methods to obtain verified, high-dependability learning systems. They include modeling a system whose structure might change and that operates in an unknown environment, generating training and testing data so as to obtain guarantees, and formulating correct-by-construction methods for designing learning systems with provable behavior. Initial encouraging results have been obtained for verifying autonomous vehicle designs that have perception components based on deep neural networks [34], a result that is possible to build upon.

Another cross-cutting challenge is to understand how the freedom (or burden) of initiative can be properly apportioned among humans and other battlefield "things" to most effectively attain mission goals and most judiciously manage risk. Allowing for initiative, lower in the command chain, is an act of delegation of the authority of decision-making. More autonomy implies less predictability of aggregate behavior which may reduce what can be guaranteed. How is that trade-off affected by other key performance requirements such as scalability, heterogeneity, and responsiveness? It may be that to attain high responsiveness and agility, or to scale to larger system sizes, more decisions need to be local, which favors autonomy, but may under some conditions impair dependability. Can systems therefore adapt the balance depending on requirements, such as acceptable response time and scale? What optimality results can be derived regarding the aforementioned trade-off and how to develop systems that are near optimal?

Needless to say, security research has a paramount role in an IoBT context. The most likely result of interconnecting battlefield entities into a big network is an increased attack surface. Novel solutions are needed to mitigate this vulnerability.

Finally, as with many other examples, where some autonomy is delegated to "things", ethical and legal implications must be addressed. Autopilots, cruise control, and collision avoidance systems are examples of today's technologies that remove humans from the decision and control loop. What are their counterparts in an IoBT context? How do they impact liability and ethics? Importantly, what decisions must remain with humans? One prime example of a human decision in a military context is the decision to fire a weapon. While we expect a human decision-maker to remain in charge of that decision in the foreseeable future, smart or learning systems can, nevertheless, improve safety of weapon use, making it more consistent with user intent. For example, smarter ammunition used in disaster response might be authorized to impact only a specific category of things (e.g., condemned buildings, such as those damaged by an earthquake). Demolition charges may use (or communicate with) sensors and computational elements to withhold from activation where humans are present, thereby reducing unintended loss of life. These and other possibilities open up when envisioning smarter connected battlefield technologies. The current IoBT project is a basic research effort that addresses the enabling intellectual foundations. Study of specific applications is deferred to future research.

VII. Conclusions

The research directions discussed above will enable a new paradigm for connecting and managing IoBT assets in army operations to meet commander's intent. In this paradigm, the system is self-aware and possesses the intelligence needed to discover and characterize new components, assemble desired mission-relevant composite assets, adapt to perturbations, recover from attacks, probe adversarial systems, monitor its own state, detect anomalies, and continuously learn from its own experiences. It aids military operations in multiple ways. First, it reduces the need for physical presence of humans in dangerous environments by improving autonomy, resilience, and survivability of assets in the field, while offering assurances on behavior. Second, the research endows mission-centric systems with adaptation reflexes to regroup and reconfigure independently, as needed to meet mission goals in response to unexpected conditions, without increasing the cognitive burden on the human operator. For example, the IoBT may independently switch to a different sensing modality in order to meet information needs of a commander upon unexpected resource losses or adverse weather conditions that render previously used modalities ineffective. Third, it explicitly supports learning over time, thereby improving with experience. The authors expect these advances to dramatically reduce the cost of conflicts and improve safety and efficacy of mission execution.

Acknowledgements

Research reported in this paper was sponsored in part by the Army Research Laboratory under Cooperative Agreement W911NF-17-2-0196. 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 Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] D. Wang, M. T. Amin, S. Li, T. Abdelzaher, L. Kaplan, S. Gu, C. Pan, H. Liu, C. C. Aggarwal, R. Ganti, X. Wang, P. Mohapatra, B. Szymanski, and H. Le, "Using humans as sensors: An estimation-theoretic perspective," in *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks (ISPN-14)*, April 2014, pp. 35–46.

[2] D. Wang, T. Abdelzaher, and L. Kaplan, *Social Sensing: Building Reliable Systems on Unreliable Data*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2015.

[3] R. W. Ouyang, M. Srivastava, A. Toniolo, and T. J. Norman, "Truth discovery in crowdsourced detection of spatial events," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 4, pp. 1047–1060, 2016.

[4] R. W. Ouyang, L. M. Kaplan, A. Toniolo, M. Srivastava, and T. J. Norman, "Parallel and streaming truth discovery in large-scale quantitative crowdsourcing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 10, pp. 2984–2997, 2016.

[5] N. Kothari, R. Gummadi, T. Millstein, and R. Govindan, "Reliable and efficient programming abstractions for wireless sensor networks," in *Proceedings of the 28th ACM SIGPLAN Conf. on Programming Language Design and Implementation*, 2007, pp. 200–210.

[6] R. Gummadi, O. Gnawali, and R. Govindan, "Macro-programming wireless sensor networks using Kairos," in *Proceedings of the International Conference on Distributed Computing in Sensor Systems*, ser. DCOSS'05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 126–140. [Online]. Available: http://dx.doi.org/10.1007/11502593_12

[7] L. Luo, T. F. Abdelzaher, T. He, and J. A. Stankovic, "EnviroSuite: An environmentally immersive programming framework for sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 5, no. 3, pp. 543–576, Aug. 2006. [Online]. Available: http://doi.acm.org/10.1145/1165780.1165782

[8] X. Gu, K. Nahrstedt, R. N. Chang, and C. Ward, "Qos-assured service composition in managed service overlay networks," in *23rd International Conference on Distributed Computing Systems, 2003. Proceedings.*, May 2003, pp. 194–201.

[9] J. Liang, X. Gu, and K. Nahrstedt, "Self-configuring information management for large-scale service overlays," in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, May 2007, pp. 472–480.

[10] B. Schlinker, R. N. Mysore, S. Smith, J. C. Mogul, A. Vahdat, M. Yu, E. Katz-Bassett, and M. Rubin, "Condor: Better topologies through declarative design," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM '15)*, 2015, pp. 449–463.

[11] Y. Chang, S. Rao, and M. Tawarmalani, "Robust validation of network designs under uncertain demands and failures," in *Proceedings of the 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)*. Boston, MA: USENIX Association, 2017, pp. 347–362. [Online]. Available: https://www.usenix.org/conference/nsdi17/technical-sessions/presentation/chang

[12] J. Heo, D. Henriksson, X. Liu, and T. Abdelzaher, "Integrating adaptive components: An emerging challenge in performance-adaptive systems and a server farm case-study," in *Proceedings of the 28th IEEE International Real-Time Systems Symposium (RTSS 2007)*, Dec 2007, pp. 227–238.

[13] J. O. Kephart, A. Diaconescu, H. Giese, A. Robertsson, T. Abdelzaher, P. Lewis, A. Filieri, L. Esterle, and S. Frey, "Self-adaptation in collective self-aware computing systems," in *Self-Aware Computing Systems*, S. Kounev, J. Kephart, A. Milenkoski, and X. Zhu, Eds. Cham: Springer International Publishing, 2017, pp. 401–435. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-47474-8_13

[14] H. V. Nguyen, R. Rivas, and K. Nahrstedt, "idsrt: Integrated dynamic soft real-time architecture for critical infrastructure data delivery over wlan," *Mob. Netw. Appl.*, vol. 16, no. 1, pp. 96–108, Feb. 2011. [Online]. Available: http://dx.doi.org/10.1007/s11036-010-0250-x

[15] L. Hoffman and N. Maier, "Quality and acceptance of problem solutions by members of homogeneous and heterogeneous groups," *Journal of Abnormal and Social Psychology*, vol. 62, pp. 401–407, 1961.

[16] L. R. Hoffman, "The group problem-solving process," in *Group Processes*, L. Berkowitz, Ed. New York: Academic Press, 1978, pp. 101–114.

[17] S. Jackson, "Team composition in organizations," in *Group Process and Productivity*, S. Worchel, W. Wood, and J. Simpson, Eds. London: Sage, 1992, pp. 1–12.

[18] C. Nemeth, "Differential contributions of majority and minority influence," *Psychological Review*, vol. 93, pp. 23–32, 1986.

[19] T. Bu, N. Duffield, F. L. Presti, and D. Towsley, "Network tomography on general topologies," *ACM SIGMETRICS Performance Evaluation Review*, vol. 30, no. 1, pp. 21–30, 2002.

[20] L. Ma, T. He, K. K. Leung, A. Swami, and D. Towsley, "Monitor placement for maximal identifiability in network tomography," in *IEEE INFOCOM 2014: Proceedings of the IEEE Conference on Computer Communications*, April 2014, pp. 1447–1455.

[21] L. Ma, T. He, A. Swami, D. Towsley, K. K. Leung, and J. Lowe, "Node failure localization via network tomography," in *Proceedings of the 2014 Internet Measurement Conference, IMC 2014, Vancouver, BC, Canada, November 5-7, 2014*, 2014, pp. 195–208.

[22] L. Ma, T. He, K. K. Leung, D. Towsley, and A. Swami, "Efficient identification of additive link metrics via network tomography," in *Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2013, pp. 581–590.

[23] P. Giridhar, M. T. Amin, T. Abdelzaher, D. Wang, L. Kaplan, J. George, and R. Ganti, "Clarisense+: An enhanced traffic anomaly explanation service using social network feeds," *Pervasive and Mobile Computing*, vol. 33, pp. 140–155, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119216000444

[24] X. Meng, J. Bradley, B. Yavuz, E. Sparks, S. Venkataraman, D. Liu, J. Freeman, D. B. Tsai, M. Amde, S. Owen, D. Xin, R. Xin, M. J. Franklin, R. Zadeh, M. Zaharia, and A. Talwalkar, "MLlib: Machine learning in Apache Spark," *Journal of Machine Learning Research*, vol. 17, no. 34, pp. 1–7, 2016.

[25] P. Carbone, S. Ewen, S. Haridi, A. Katsifodimos, V. Markl, and K. Tzoumas, "Apache Flink$^{\text{TM}}$: Stream and batch processing in a single engine," *Bulletin of the Technical Committee on Data Engineering*, vol. 38, no. 4, pp. 28–38, December 2015.

[26] Z. Kang, K. Grauman, and F. Sha, "Learning with whom to share in multitask feature learning," in *Proceedings of the 28th International Conference on Machine Learning (ICML)*, Bellevue, WA, 2011, pp. 521–528, (acceptance rate: 26%, citations: 25).

[27] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," *Computing Research Repository (CoRR)*, vol. abs/1412.6572, 2014.

[28] R. Kolte, A. Ozgur, and S. N. Diggavi, "When are dynamic relaying strategies necessary in half-duplex wireless networks?" *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1720–1738, 2015.

[29] C. Karakus, I.-H. Wang, and S. N. Diggavi, "Gaussian interference channel with intermittent feedback," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4663–4699, 2015.

[30] S. Mishra, I.-H. Wang, and S. N. Diggavi, "Harnessing bursty interference in multicarrier systems with output feedback," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4430–4452, 2017.

[31] L. Keller, M. Siavoshani, C. Fragouli, K. J. Argyraki, and S. N. Diggavi, "Joint identity-message coding," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1083–1093, 2010.

[32] J. Sebastian, A. Sengupta, and S. N. Diggavi, "On capacity of noncoherent mimo with asymmetric link strengths," in *IEEE International Symposium on Information Theory (ISIT)*, June 2017.

[33] J. Sebastian and S. N. Diggavi, "Non-coherent parallel relay networks," 2017, available on Arxiv.

[34] T. Dreossi, A. Donzé, and S. A. Seshia, "Compositional falsification of cyber-physical systems with machine learning components," in *Proceedings of the NASA Formal Methods Symposium (NFM)*, ser. Lecture Notes in Computer Science, vol. 10227. Springer, May 2017, pp. 357–372.

[35] S. Dutta, S. Jha, S. Sanakaranarayanan, and A. Tiwari, "Output range analysis for deep neural networks," *CoRR*, vol. abs/1709.09130, 2017. [Online]. Available: http://arxiv.org/abs/1709.09130

[36] S. A. Seshia, "Combining induction, deduction, and structure for verification and synthesis," *Proceedings of the IEEE*, vol. 103, no. 11, pp. 2036–2051, 2015.

[37] S. A. Seshia, D. Sadigh, and S. S. Sastry, "Towards Verified Artificial Intelligence," *ArXiv e-prints*, July 2016.