AI at the Tactical Edge

Alexia Schulz and Pierre Trepagnier
Cyber Physical Systems
MIT Lincoln Laboratory
244 Wood St.
Lexington, MA 02421

***Abstract*—Artificial intelligence (AI) and Machine Learning (ML) are assuming an increasingly prominent role in the modern armamentarium of the warfighter, ideally acting as an always-on-duty assistant. In this Extended Abstract we explore aspects of AI/ML which are particularly characteristic of its deployment at the tactical edge, by which we mean warfighters directly involved in executing the mission at the "tip of the spear [1]." AI intrinsically depends on compute power and communications. At the tactical edge both these resources are generally in short supply, expensive to provision, and shared among contending needs at the best of times, let alone in more critical situations. Here we enumerate a number of possible applications of AI/ML at the tactical edge, characterizing them by features such as compute power and data required, both at train time and run time. From these illustrative examples we generalize a set of suitability characteristics for tactical edge AI applications to best ensure that they contribute to warfighter resilience rather than fail at the time of greatest need.**

***Index Terms*—Artificial Intelligence, Tactical Edge**

I. Introduction

A. Organization and Resources of the Tactical Edge

The tactical edge is considered to be the "tip of the spear" in mission execution, however the actual organization of warfighters at the tactical edge is complicated and varies from service to service. For our purposes, we will adopt a simplified and stylized three-tier edge model based on compute resources:

Tier 1: ("Squad") This lowest tier consists of warfighters who have access to handheld devices and perhaps additional compute capacity on the order of a laptop. The laptop, however, may have multiple contending roles. Communication is limited to tactical radios.

Tier 2: ("Battalion") We group this tier with Tier 1 as part of the edge, but it will generally contain several hundred to a thousand warfighters. Assume at least a 5U rack of servers are present, but under severe resource contention. Communications will consist of SATCOM [2] as well as tactical radio; WiFi will generally be present within the battalion headquarters. Communication between Tiers 2 and 1 is spotty, and may have to be effected by sneakernet in case of adversarial jamming.

Tier 3: ("Center") This tier is "everything not at the edge" and is assumed to have abundant computational and data storage resources. However, communication between Tier 3 and the other two tiers is intermittent and bandwidth constrained in the best of times. In the case of near-peer adversarial activity it should be assumed to be nonexistent [3].

The air and sea analogs to each of these tiers exist. For example an individual fighter might be Tier 1 whereas an ISR asset better tracks to Tier 2. A ship is most often at Tier 2.

B. AI/ML Characteristics Relevant to the Tactical Edge

AI/ML is an enormous and active area of research, and enumerating their characteristics is well beyond the scope of this short paper. Here we shall confine ourselves to simply enumerating those characteristics which are of particular importance to their use at the tactical edge. The detailed interplay of these characteristics and the tactical edge will be the subject of the next section.

- AI/ML models have an inherent time constant - a δT over which they remain valid before they go stale. However this δT varies with the model domain.
- AI/ML models require disparate compute resources to generate, depending on the task and the data. Compute resources are typically higher in the center than at the edge.
- Model training is more compute resource intensive if data volume is higher, if model parameters are many, or both.
- Because of constrained resources, model training at the edge is only practical for small data volumes and few model parameters.
- However, sometimes a model can be trained with large computing resources, and later tuned with small resources. This process is referred to as transfer learning, and has been applied successfully in fields such as low-resource languages [4]. Transfer learning works when the new and old models are fundamentally similar, and the model representation is stable to the addition of new training data.

II. Exemplary Tactical Edge AI Applications

Having discussed briefly the characteristics of the tactical edge and how they interact with the relevant parts of AI/ML, we now turn to specific examples which serve to illustrate the general points. These are listed in Table 1. The rows in the table are specific examples of AI/ML at the tactical edge, either potential or actual. The columns are AI/ML characteristics which determine suitability for deployment, as discussed above. Let us consider them in a little more detail.

Train Time Characteristics: This is broken into three parts which together determine required computing resources

Exemplary Applications of AI at Tactical Edge	Train time			Run time		Model stability	Training Data		Run-Time	Suitability for Tactical Edge	
	Volume of data	Number of model parameters	Transfer learning or full retrain	Volume of data required	Compute power required	Long or short compared to assumed comms degradation	Acquistion	Processing	Data Processing	Unsophisticated Adversary	Near-Peer Adversary
Navigation and UAX											
Map navigation instructions for humans	medium	low	transfer	low	low	short	center	center	edge		
GPS guided navigation for UAX	high	medium	transfer	medium	medium	long	center	center	edge		
Navigation in GPS denied environment	high	medium	transfer	medium	medium	long	both	center	edge		
Surveillance											
Personnel ID through devices	medium	small	full	medium	low	medium	edge	center	edge		
Peresonnel ID through facial recognition/gait analysis	high	medium	full	high	medium/high	long	edge	center	edge		
Identification of adversary comms	medium	medium	full	medium	low	long(U) short(NP)	both	center	edge		
Automated processing of video surveillance	high	high	full	high	high	short	both	center	center		
Targeting and Maneuver											
Target selection or identification	medium	medium	full	low	medium	long	both	center	edge		
Target tracking in a complex environment	medium	high	transfer	low	low	short	edge	center	edge		
Sensor-to-shooter fires	low	medium	full	low	low	long	edge	center	edge		
Target spoofing for defensive purposes	high	high	transfer	low	low	long	center	center	edge		
Aircraft Operations and Maintenance											
Fly-by-wire	medium	medium	full	small	low	long	center	center	edge		
Maintenance customization via data analysis	medium	medium	full	medium	medium	long	center	center	center		
Language Applications											
Speaker ID (voice, handwriting, etc.)	medium	high	full	medium	medium	long	edge	center	edge		
Speech recognition and transcription	high	high	full	high	medium	long	edge	center	edge		
Automated translation	high	high	full	low	medium	long	edge	center	edge		
Network configuration											
ID of compromised assets and/or automated quarantine	medium	medium	full	high	high	medium	center	center	center		
Filtering and prioritization of vulnerabilities	medium	medium	full	medium	low	short	center	center	edge		
Mobile ad-hoc network node distribution and coverage	low	medium	transfer	low	low	short	edge	edge	edge		
OCO/EW											
False targets, deepfakes, and/or subterfuge	high	high	full	high	medium	long	both	center	edge		
Battle damage assessment	high	medium	full	high	medium	long	both	center	both		
Delivery of cyber effects	medium	medium	full	medium	medium	long	center	center	edge		
Optimal jamming strategy	medium	low	transfer	medium	medium	long	center	center	edge		
Miscellaneous											
Battlefield medical triage	high	high	full	low	low	long	center	center	edge		
Automated sitrep report generation	high	high	full	medium	medium	long	center	center	edge		
Noise rejection/suppression in battlefield comms	medium	medium	full	medium	low	long	center	center	edge		

Fig. 1: The suitability of different AI/ML applications at the tactical edge. Rows indicate different AI/ML exemplars, and columns indicate factors that render these algorithms robust and resilient for warfighters relying upon them. Communications external to the edge are assumed to be intermittent for an unsophisticated adversary, and non-existent for a near-peer.

- volume of data, number of model parameters, and whether or not transfer learning is possible.

Run Time Characteristics: These are divided into volume of data required, and compute power required at run time.

Model Stability: Communication with the center should be assumed to be severely degraded or (in the case of a near-peer power) blocked altogether during conflicts. This column classifies model stability on timescales relative to the communication disruption.

Training Data: These two columns specify whether the train-time data is acquired or processed at the edge or center. For models with short δT, training at the edge is preferable.

Run Time: This specifies whether at run time, communication with the center is required or if the AI/ML can run autonomously at the edge.

Suitability: Finally, at the right, the information is rolled up into a stoplight chart suggesting the overall suitability of the example AI/ML. Note it is subdivided, the difference being that an unsophisticated adversary is assumed to be able to degrade communications, while a near-peer one is assumed to be able to block them completely.

III. Conclusion

The most restrictive and safest takeaway from Table 1 is that if warfighters at the edge are to be confident that they have AIs they can depend on in times of conflict, they require AIs running on equipment also at the edge and under their direct control. Maximal resilience at the edge depends on not attempting to use AI/ML systems that require significant compute power or any but the most minimal communications. Cloud computing, for instance, is out of the question.

However there are additional subtleties having to do with model training and stability as well as compute power worth exploring. A stable model can require large resources to train without becoming problematic if it can be run at the edge with available data and compute power, since the trained model can be transferred from the center to the edge prior to conflict. (Note, however, sophisticated actors that deploy previously unseen Wartime Reserve Modes [5] may generate incorrect predictions from the model). Similarly, even relatively unstable models which require training at the center could be of use, so long as transfer learning can be employed. To summarize, in order to promote resilience at the tactical edge, we advocate careful selection of AI/ML systems based on the factors of minimizing required compute power, and data volume, while maximizing model stability.

Acknowledgments

Thanks to LtCol Francisco Caceres and also the members of our AI at the Tactical Edge focus group: Gregory Angelides, Jonas Bengt, Kevin Brady, Joseph Campbell, Michael Coury, Jeffrey Gottschalk, Jacob Harer, Jarilyn Hernandez Jimenez, Nicolas Malyska, Jeremy Mineweaser, Zachary Ravichandran, Sage Trudeau, Allan Wollaber, and Joseph Zipkin.

This material is based upon work supported by the United States Air Force under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force.

REFERENCES

[1] F. Dandashi, P. Glasow, D. Kaplan, W. Lin, S. Semy, J. Valentine, and B. Yost, "Tactical edge characterization framework," 2007.

[2] PEO-C3T, "Satellite communications t2c2," *https://peoc3t.army.mil/tn/t2c2.php*, 2020.

[3] C. Kawasaki, "Four future trends in tactical network modernization," *https://www.army.mil/article/216031/*, 2019.

[4] B. Zoph, D. Yuret, J. May, and K. Knight, "Transfer learning for low-resource neural machine translation," *arXiv preprint arXiv:1604.02201*, 2016.

[5] W. E. Gortney, "Dictionary of military and associated terms," 2012.