Resilient Cooperative Control of Cyber Physical Systems

Xenofon Koutsoukos and Heath LeBlanc
Institute for Software Integrated Systems (ISIS)
Vanderbilt University
Xenofon.Koutsoukos@vanderbilt.edu, heath.j.leblanc@vanderbilt.edu

Introduction

The advance of embedded systems and networking technology has facilitated a paradigm shift in engineering system design, from centralized to distributed. This shift has lead to significant interest in the design and analysis of *multi-agent networks*. A multi-agent network, or *networked multi-agent system*, consists of a set of *agents*, or *nodes* that may represent processors, robots, and so on. The agents are equipped with *sensing* and/or *communicating*, along with computational resources and possibly *actuation*. Through a *network*, the agents share information in order to achieve specific *group objectives*. Examples of group objectives include consensus, synchronization, formation control, and cooperative load transport. In order for the group objectives to be achieved, *distributed algorithms* are used to coordinate the behavior of the agents.

One of the fundamental challenges in the design of networked multi-agent systems is that the coordination algorithms use only *local information*, i.e., information obtained by the individual agent through sensor measurements, calculations, or communication with neighbors in the network. Another challenge lies in the fact that not only is each agent a dynamical system, but the network itself is dynamic. Therefore, the distributed algorithms must be designed to handle time-varying network topologies. Information may not be able to be relayed across the dynamic network in a reliable manner. A third challenge is caused by uncertainties introduced by the network and in the implementation of the control and coordination algorithms.

More importantly, multi-agent networks, like all large-scale distributed systems, have many entry points for malicious attacks or intrusions. If one or more of the agents are compromised in a security breach, it is crucial for the networked system to continue operating with minimal degradation in performance and the success of the global objective should be assured. To achieve this, it is necessary for the cooperative algorithms to be designed in such a way that they can withstand the compromise of a subset of the nodes and *still guarantee some notion of correct behavior at a minimum level of performance*. We refer to such a multi-agent network as being *resilient* to adversaries.

This summary describes a framework for resilient cooperative control in networked multi-agent systems using low complexity distributed algorithms for solving consensus and synchronization problems in dynamical systems. The framework combines methods from distributed computing and control engineering to devise coordination algorithms that ensure that group objectives such as consensus and synchronization are satisfied even in the presence of adversaries. The work can contribute to providing scalable computational methods for resilient cooperative control in the presence of adversaries while ensuring robustness to real-world system behavior and interactions in dynamic networks.

Resilient Consensus and Synchronization

Reaching consensus is fundamental to group coordination, and involves reaching agreement on a quantity of interest. There are several variations of how consensus problems are defined and can be employed for cooperative control. Synchronization in networked multi-agent systems is defined as a dynamic form of agreement on all of the state (or output) variables of the agents. From one perspective, it is a strict generalization of consensus, where the agents have more complicated dynamics than simply an integrator or accumulator. Synchronization problems tend to deal exclusively with agreement on physically dependent quantities. Therefore, synchronization can be viewed as a strict generalization of physically dependent consensus, and entails agreement whenever the states are oscillatory or converge to limit cycles.

There is a long history in distributed computing of studying consensus problems in the presence of faults and adversarial processors. The most potentially harmful form of adversary is the Byzantine processor, which may behave arbitrarily within the limitations set by the model of computation. Typically, the number of processors that may be Byzantine are bounded and fundamental tight bounds have been established on the ratio of Byzantine to normal processors, as well as on the connectivity of the network.

From a control theoretic viewpoint, consensus and synchronization in the presence of adversaries has only been considered recently, and has focused on detection and identification of misbehaving agents in linear consensus networks. While detection is clearly an important problem, these techniques require each agent to have information of the network topology beyond its local neighborhood. This requirement of *nonlocal information* renders these techniques inapplicable to general time-varying networks. Further, the detection algorithms do not consider safety constraints on the

states of the agents and it is possible that the adversaries may compromise the networked system before detection, which may not be suitable for certain safety critical applications.

We have investigated coordination protocols to achieve resilience against adversaries in network multi-agent systems [1], [2], [3]. We consider both time-invariant and time-varying (or switching) network topologies. The protocols have low complexity, use only local information, and they do not require historical information. Our analysis shows that traditional graph theoretic metrics are inadequate to characterize the tight topological conditions under which convergence is assured. Indeed, any resilient distributed algorithm that is capable of succeeding in the presence of adversaries, without a priori knowledge of the adversaries or network topology, must filter the information from neighboring nodes with some measure of skepticism. This amounts to the need for redundancy of information from neighbors in the network, and is the basis of the novel property referred to as *network robustness*.

The adversaries model compromised nodes that are hijacked in a security breach. The model for the adversaries consists of a threat model and a scope of threat assumption. The threat model defines the types of behaviors allowed by the adversaries. We study Byzantine, malicious, and crash threat models. A Byzantine adversary is similar to a Byzantine fault; it may behave in an arbitrary fashion (within the scope of the model of computation) and may convey different information to different neighbors in the network. The malicious adversary is similar to the Byzantine model, but malicious nodes must convey the same information to all neighbors (i.e., the malicious adversary is a local broadcast version of the Byzantine adversary). Finally, the crash adversary may select a time to crash the compromised node, which forces the states of the compromised node to remain constant. All adversaries are considered omniscient. This means they have access to global information concerning the multi-agent network, including the full network topology, the algorithms used by the other nodes, the states of the other nodes, which nodes are compromised, and the plans of the other adversary nodes. One may take the viewpoint that the individual adversary nodes are controlled by a central commanding adversary.

The scope of threat assumptions bound the number or fraction of compromised nodes either in the entire network or in the neighborhood of any normal node. Whenever, the total number of compromised nodes in the network is assumed to be bounded above by some constant $F \in \mathbb{Z}_{\geq 0}$, we say the adversaries satisfy the F-total model. In cases where it is preferable to make *no global assumptions*, we are interested in other threat assumptions that are strictly local. For example, whenever each node assumes that at most F nodes in its *neighborhood* are compromised (but there is no other bound on the total number of compromised nodes), the scope of threat is F*-local*. Alternatively, if it is assumed that there is an upper bound on the *fraction*, $0 \leq f \leq 1$, of compromised nodes in any normal node's neighborhood, we say the adversaries satisfy the f-fraction local model.

The goal is for the uncompromised nodes, or simply normal nodes, to still achieve the group objective in the presence of the adversary nodes. Therefore, the networked multi-agent system should be resilient to adversaries. It is important to emphasize that the only type of security breach we study is compromised (adversary) nodes, as opposed to attacks on the communication network (e.g., denial of service or deception attacks). That being said, the models considered are general enough so that from a local perspective (i.e., from the point of view of the individuals in the network), the difference between compromising the node or the outgoing communications of the same node is indistinguishable.

We have derived theoretical conditions in terms of network robustness for solving the adversarial consensus and synchronization problems (1) under different timing models including continuous time, synchronous discrete time, and asynchronous discrete time, (2) under different adversary models including crash, malicious, and Byzantine agents, and (3) under both global and local assumptions on the scope of the threats.

Future Work

There are several interesting directions for future work. For example, the techniques for resilience can be extended to more complex network architectures that include hierarchy and consist of heterogeneous agents. Further, the agents can be considered to be mobile, which leads to an important co-dependence between the locality of the agents and the topology of the network. Examining ways to maintain network robustness in the presence of adversaries under spatial constraints is an interesting and important research problem.

Acknowledgments

This work is supported in part by the National Science Foundation (CNS-1035655, CCF-0820088), the U.S. Army Research Office (ARO W911NF-10-1-0005), and Lockheed Martin.

References

[1] LeBlanc, H. J. *Resilient Cooperative Control of Networked Multi-Agent Systems*. PhD thesis, Department of Electrical Engineering and Computer Science, Vanderbilt University, 2012.

[2] LeBlanc, H. J., and Koutsoukos, X. D. Consensus in networked multi-agent systems with adversaries. In *Proceedings of the 14th international conference on Hybrid systems: computation and control* (Chicago, IL, 2011), (HSCC '11), pp. 281–290.

[3] LeBlanc, H. J., and Koutsoukos, X. D. Low complexity resilient consensus in networked multi-agent systems with adversaries. In *Proceedings of the 15th international conference on Hybrid systems: computation and control* (Beijing, China, 2012), (HSCC '12), pp. 5–14.