Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Resilient Consensus against Mobile Malicious Agents

Yuan Wang, Hideaki Ishii, François Bonnet, Xavier Défago *

* *Dept of Computer Science, Tokyo Institute of Technology, Japan*
wang@sc.dis.titech.ac.jp, {ishii, bonnet, defago}@c.titech.ac.jp

Abstract: This paper addresses consensus problems in the presence of adversaries that can move within the network and induce faulty behaviors in the attacked agents. By employing mobile adversary models from the computer science literature, we develop three protocols which can mitigate the influence of malicious agents. The algorithms follow the class of mean subsequence reduced (MSR) algorithms, under which agents ignore the suspicious values received from neighbors during their state updates. Different from the static model, even after the adversaries move away, the infected agents may remain faulty in their values for a short while, which must be taken into account. We develop conditions on the network structures for both the complete and non-complete graph cases, under which the proposed algorithms are guaranteed to attain resilient consensus. An illustrative example is provided to verify the effectiveness of our approach.

Keywords: Distributed algorithms, Resilient multi-agent consensus, Mobile adversary agents.

1. INTRODUCTION

In the domain of cyber-physical systems, security problems have recently become a critical issue. Cyber attacks can cause damages not only from having important information stolen, but also from having physical equipments manipulated, which may lead to serious faults and accidents. Security related problems have been investigated in a range of disciplines including computer science, control, robotics, and power systems (Sandberg et al. (2015)).

In this paper, we follow the research on fault-tolerant distributed algorithms and focus on resilient consensus problems. Consensus forms a fundamental problem in multi-agents systems (Lynch (1996); Mesbahi and Egerstedt (2010)), where agents communicate with each other to reach the global objective to share a common value. In uncertain environments, adversaries may attack other agents to change their behaviors, which can potentially keep the regular agents from reaching consensus. Hence, it is important to guarantee such regular agents to remain resilient from adversarial attacks.

In particular, we deal with adversaries that can switch the target agents from time to time. Such mobile adversaries can cooperate in a worst-case manner by communicating with each other even if no direct link is present. On the other hand, the attacked agents may recover and become fault-free again though the agent's value may still be corrupted. Depending on the awareness of the agent itself, it can take different actions. For example, it can use only the neighbors' values for starting new in the consensus process. Such recovery may be performed by reboot or reset of the system manually by the system operator or automatically by devices such as watchdogs.

⋆ This work was supported in the part by the JST CREST Grant No. JPMJCR15K3 and by JSPS under Grant-in-Aid for Scientific Research Grant No. 18H01460.

For mobile adversaries, several models have been proposed in the literature (Buhrman et al. (1995); Garay (1994); Sasaki et al. (2013); Bonnet et al. (2016)). They are different in terms of the timings of attacks for the adversaries and the capabilities of the agents recovering from infections. Recently, by Bonomi et al. (2019), these studies have been extended to the case where the agents' states take real values. However, all of these studies have been limited to networks taking complete graph forms and moreover to Byzantine adversaries. Such adversaries are the worst type as they can freely manipulate their states and are capable to send different messages to their neighbors.

The contribution of this work is threefold: First, we extend the mobile adversary model in the real-valued states case to the socalled malicious adversary models. Malicious agents form a subclass of the Byzantine in that they can only broadcast data, that is, they send the same data to all neighbors. Second, we propose novel protocols for achieving resilient consensus under three different mobile malicious models. The protocols follow the resilient approach known as the mean subsequence reduced (MSR) algorithms (Kieckhafer and Azadmanesh (1994)). In updating their state values, the agents ignore the suspicious values sent by other agents. Third, we consider networks taking non-complete graph forms and characterize the necessary connectivity structures for the proposed MSR-based protocols to guarantee resilient consensus.

The considered problem setting is natural from the viewpoint of applications such as wireless sensor networks, where agents communicate with a limited number of neighboring agents and use broadcast transmissions. Moreover, our results have been motivated by the recent advances made in resilient consensus problems initiated by LeBlanc et al. (2013) and Vaidya et al. (2012). There, for MSR algorithms, tight characterizations on the network structures have been made by introducing the notion of graph robustness. Further works can be found in, e.g., (Chen

et al. (2018); Dibaji and Ishii (2015); Dibaji et al. (2018); Wang and Ishii (2020, 2019); Zhang et al. (2015)).

Concerning mobile adversary agents, the early work by Buhrman et al. (1995) has proposed a model where the malicious agents can move and switch their identities; when they move away, the recovering agents are cured from infections immediately and can be treated as regular in the next step. Another work by Garay (1994) discusses a more general model where the cured agent can detect the infection when they recover. Recently, other mobile adversary models and resilient algorithms have been proposed by Sasaki et al. (2013) and Bonnet et al. (2016), where the detection of infection by the cured agents is not possible. We extend these models to agents whose states take real values under the malicious adversary model. In this case, in fact, the two models in (Sasaki et al. (2013)) and (Bonnet et al. (2016)) coincide. When the adversaries are mobile, the conventional MSR-based algorithms for the static adversaries mentioned above cannot guarantee resilient consensus. This is mainly because the recovering agents require special attention as they may have corrupted values.

This paper is organized as follows. In Section 2, the resilient consensus problem is formulated. We propose three protocols for three mobile malicious models in Sections 3 to 5. In our analysis, we provide sufficient conditions on required network structures. An example is provided in Section 6 to illustrate the effectiveness of the algorithms. We conclude this work in Section 7. Proofs of the results are omitted for space reasons.

2. PROBLEM FORMULATION

2.1 Graph Notions

Denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ a graph consisting of n nodes, where the set of nodes is $\mathcal{V} = \{1, 2, \ldots, n\}$ and the set of edges is $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The edge $(j, i) \in \mathcal{E}$ indicates that node j can send a message to node i and is called an incoming edge of node i. Directed graphs are considered, in which $(j, i) \in \mathcal{E}$ does not necessarily imply $(i, j) \in \mathcal{E}$. Let $\mathcal{N}_i = \{j : (j, i) \in \mathcal{E}\}$ be the set of (in-)neighbors of node i. The path from node i_1 to node i_p is denoted as the sequence $(i_1, i_2, \ldots, i_p)$, where $(i_j, i_{j+1}) \in \mathcal{E}$ for $j = 1, \ldots, p-1$. The graph $\mathcal{G}$ is said to have a spanning tree if there exists a node from which there are paths to all other nodes in this graph. Moreover, the graph is said to be complete, if for each pair of nodes, there is a bidirectional edge $(i, j) \in \mathcal{E}$ connecting them.

2.2 Mobile Malicious Agents and Resilient Consensus

We consider a multi-agent system with n agents interacting over the directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Each node has a state $x_i(k)$, which takes a real value. The objective of consensus is that starting from initial values $x_i(0)$, all agents update their states iteratively by communicating with their neighbors to arrive at the same value as $\lim_{k\to\infty} |x_i(k) - x_j(k)| = 0$ for $i, j \in \mathcal{V}$.

In this paper, we study multi-agent systems situated in an uncertain environment, where some agents are faulty and/or adversarial and do not execute the given algorithm properly by updating their states arbitrarily. We introduce a new class for such faulty agents, which is called *mobile malicious model*. Informally, this class has the following two features: (i) The adversary agents may transmit their false states to their neighbors through broadcast, i.e., all neighbors of a malicious agent receive the same data. (ii) The identity of the malicious agents can switch over time. That is, an attacker may turn a non-adversarial agent to become malicious at some time. Also, a malicious agent may recover and be *cured* at a later time. It is said to be mobile to indicate that the attacker may switch between different agents in infecting them. In this work, we treat the mobile agents deterministically though they share similarities with stochastic models studied for spreading processes of infectious diseases (Nowzari et al. (2016)).

We provide more notations and notions for the mobile models considered in this paper. At each time k, the set $\mathcal{V}$ of nodes is partitioned into two subsets: The set $\mathcal{R}(k)$ of regular agents and the set $\mathcal{A}(k)$ of adversarial agents. In the static case, $\mathcal{R}(k)$ and $\mathcal{A}(k)$ remain invariant over time. Their faulty and abnormal behaviors are defined below.

Definition 1. (*Malicious*): An adversarial agent $i \in \mathcal{A}(k)$ is said to be malicious if it makes updates in its value $x_i(k)$ arbitrarily and sends the same value to all of its neighbors each time a transmission is made.

The notion of malicious agents is natural in many applications. For example, in wireless sensor networks, each sensor node communicates by broadcasting its data, and hence its neighbors receive the same state data.

Different from the static version of malicious models studied in, e.g., (LeBlanc et al. (2013)), the mobile adversaries can exhibit more variety in their behaviors. As discussed later, we will adopt three classes of such mobile adversary models from the literature, which consider mainly the Byzantine agents (Bonomi et al. (2019)). Under the mobile adversary model, the identity of the adversaries may switch, but we limit their influence by bounding the total number of them in the network over time. This is called the F-total model as defined below.

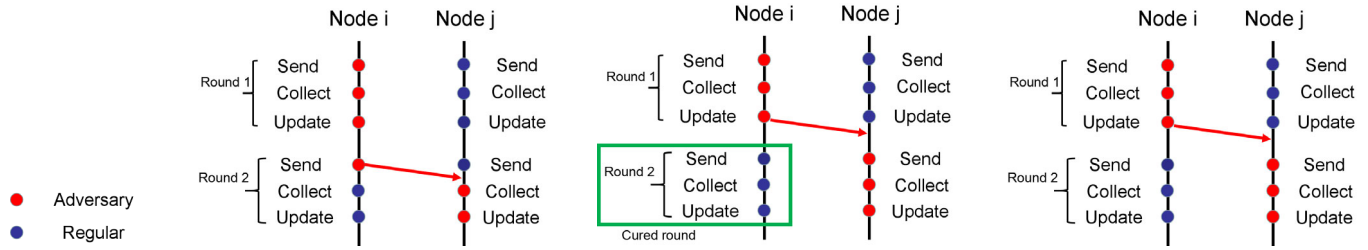
Definition 2. (F-*total*): The mobile adversarial set $\mathcal{A}(k)$ follows the F-*total* model if $|\mathcal{A}(k)| \leq F$ for all k.

For the multi-agent system in the presence of mobile adversary agents, we provide the notion of resilient consensus. Denote the maximum and minimum values among the states of the regular agents by $\overline{x}(k) = \max\{x_i(k) : i \in \mathcal{R}(k)\}$, $\underline{x}(k) = \min\{x_i(k) : i \in \mathcal{R}(k)\}$, respectively. These are the values that should eventually become the same in our resilient setting. Note that these values are chosen only among the regular agents in $\mathcal{R}(k)$ at time k.

Definition 3. (*Resilient consensus*): If for any possible sets and behaviors of the mobile malicious agents and any initial states of the regular agents, the following conditions are satisfied, then the resilient consensus is reached:

1. Safety condition: There exists an interval $\mathcal{S} \subset \mathbb{R}$ such that $x_i(k) \in \mathcal{S}$ for all $i \in \mathcal{R}(k)$, $k \in \mathbb{Z}_+$.
2. Consensus condition: The regular agents eventually take the same value as $\lim_{k\to\infty} \overline{x}(k) - \underline{x}(k) = 0$.

This paper aims to develop distributed algorithms for the regular agents in the system to reach resilient consensus.



(a) Nodes (b) Buhrman's mobile model (M1) (c) Garay's mobile model (M2) (d) Bonnet's mobile model (M3)

Fig. 1. Mobile adversary models for the malicious agents case

This problem is an extension of those studied in (Bouzid et al. (2010); LeBlanc et al. (2013); Dibaji and Ishii (2015)), which are limited to the static adversary models. Under the mobile adversary model, the notion of resilient consensus is slightly different from the static case. The agents in the adversary status, and in some cases, those in the recovering status, need not be in consensus with others. As a consequence of the attacks and/or faults, the adversary and recovering agents can take arbitrary values even outside the safety interval $\mathcal{S}$.

To mitigate the influence of the adversaries, we develop modified versions of the socalled *mean subsequence reduced (MSR)* algorithms. For the static case, such algorithms are capable to realize resilient consensus. For each regular agent i, the update rule of its state $x_i(k)$ is written as

$$x_i(k+1) = x_i(k) + \sum_{j \in \mathcal{M}_i(k)} a_{ij}(k)\left(x_j(k) - x_i(k)\right), \quad (1)$$

where the weights must satisfy $a_{ij} \in [\gamma, 1)$ with $\gamma \in (0, 1/2)$ and $\sum_{j\in\mathcal{M}_i(k)} a_{ij}(k) \leq 1$. Here, $\mathcal{M}_i(k)$ denotes the subset of agent i's neighbor set $\mathcal{N}_i$, whose states do not take extreme values; informally, among the neighbors, the F largest and the F smallest values are removed to mitigate the influence of the malicious agents. It is known that to guarantee resilient consensus by the MSR algorithm under the F-total model, it is necessary and sufficient that the network topology satisfies a condition expressed in terms of its connectivity. More specifically, the network must have a property known as $(F+1, F+1)$-robustness; see, e.g., (LeBlanc et al. (2013)).

However, we can show that mobile adversaries can easily destroy resilient consensus if the conventional approach for the static F-total model is directly applied. One issue is related to the presence of recovering nodes. Suppose that, at one time, the adversary moves to a different normal agent, which becomes malicious. In the meantime, the agent which was infected now recovers and becomes normal. Such a recovering node might have a corrupted value left from the attack. At this moment, there are more agents taking abnormal values in the network even if the attacker is capable to infect only one agent at a time. In such circumstances, the conventional MSR algorithms for the F-total model cannot guarantee resilient consensus.

In our analysis, it is more convenient to use an alternative expression of the update rule (1). Let the self-weight $a_{ii}(k) = 1 - \sum_{j\in\mathcal{M}_i(k)} a_{ij}(k)$ and the extended neighbor set $\mathcal{M}_i^+(k) = \{i\} \cup \mathcal{M}_i(k)$ containing the index of node i itself. Then, we can rewrite (1) as $x_i(k+1) = \sum_{j\in\mathcal{M}_i^+(k)} a_{ij}(k)x_j(k)$.

2.3 Models for Mobile Malicious Behaviors

Here, we introduce three classes of mobile malicious behaviors denoted as models M1, M2, and M3. The differences are related to what happens when an adversary moves to another agent and, especially, to whether the recovering agent is aware that it was attacked and its state may be corrupted during the attack. These models are taken from the literature in computer science originally developed for Byzantine adversaries. We propose versions adapted for the malicious adversaries case below. The models are illustrated in Figs. 1(b)–1(d).

It is noted that for the state update in the MSR algorithm, each agent executes three basic steps (LeBlanc et al. (2013)): *Send*, *collect*, and *update*. At time (or *round*) k, first, a regular agent i broadcasts its current value $x_i(k)$ to its neighboring agents. Second, it collects the values of the neighbors $x_j(k)$ for $j \in \mathcal{N}_i$. Third, after deleting some of the neighbor values, the value is updated to $x_i(k+1)$.

- M1 Buhrman's model (Buhrman et al. (1995)): The adversary can move away from an attacked agent i only at the sending step in each round k (Fig. 1(b)). This means that agent i broadcasts its corrupted state $x_i(k)$ to neighbors, but then becomes recovered immediately. Hence, agent i can collect and update its state as a regular node. For this reason, agent i will be classified as regular in the this round k, i.e., $i \in \mathcal{R}(k)$. If the adversary moved from agent i to another agent j after the send step, then we have $j \in \mathcal{A}(k)$. It is important to note that at each round, there are at most F faulty values in the network.
- M2 Garay's model (Garay (1994)): In this model, each agent has an additional variable, the *cured flag* $\theta_i(k)$; initially, it is set as $\theta_i(k) = 0$. The adversary can move away from an attacked agent i to agent j at any step in each round k (Fig. 1(c)). In this model, agent i is classified as adversarial at round k, i.e., $i \in \mathcal{A}(k)$, and as regular in the next round $k+1$, i.e., $i \in \mathcal{R}(k+1)$. In round $k+1$, agent i is aware that it was infected and sets its flag as $\theta_i(k) = 1$. In this state, it does not send its potentially corrupted value to neighbors nor does it use its own value during its update; the flag is set back to $\theta_i(k) = 0$ after the update step in round k. At each round, there are at most F faulty values and F missing values in the network.
- M3 Bonnet's model (Bonnet et al. (2016)): As in M2 above, the adversary can move away from an attacked agent i at any step during each round k (Fig. 1(d)). Thus, we have $i \in \mathcal{A}(k)$ and $i \in \mathcal{R}(k+1)$. At round $k+1$, agent i in the cured state is however not aware that it was infected, and hence makes the

next update as usual. There are at most $2F$ faulty values in the network: F of them are due to attacks and the remaining F from cured agents like agent i.

To deal with each of these models, we provide three protocols in the following sections.

3. PROTOCOL 1 FOR THE M1 MODEL

Here, we present the first protocol for the mobile adversaries, which is a modified version of the MSR algorithm from, e.g., (LeBlanc et al. (2013); Dibaji and Ishii (2015)). It will be shown that this protocol is effective to deal with mobile malicious agents under the model M1.

Protocol 1. At each round k, regular agent $i \in \mathcal{R}(k)$ executes the following three steps:

1. *(Send)* It broadcasts its current value $x_i(k)$.
2. *(Collect)* It collects values $x_j(k)$ of neighbors $j \in \mathcal{N}_i$.
3. *(Update)* It sorts the received values and its own value in descending order. Agent i then deletes the F largest and the F smallest values, which will not be used in the update. The set of indices of agents whose values remained is written as $\mathcal{M}_i^+(k) \subset \{i\} \cup \mathcal{N}_i$. Finally, agent i updates its value by (1).

A unique feature of this algorithm is that agent i might not use its own value. This is because in Step 3, $2F$ values are deleted regardless of agent i's value. By contrast, in conventional algorithms for static adversary models in (LeBlanc et al. (2013); Dibaji and Ishii (2015)), this number depends on the current value of agent i.

We establish that with Protocol 1, we can achieve resilient consensus under the M1 model (Buhrman et al. (1995)). Here, we first present the result for networks in the complete graph form.

Proposition 1. Consider the multi-agent system whose network $\mathcal{G}$ forms a complete graph. Suppose that the mobile malicious agents follow the F-total and M1 model. Then, the regular agents using Protocol 1 reach resilient consensus if and only if $n \geq 2F + 1$. The safety interval is given by $\mathcal{S} = [\underline{x}(0), \overline{x}(0)]$.

This proposition can be seen as an extension of a result given in (Bonomi et al. (2019)), which deals with Byzantine-type mobile adversaries. The condition there is $n \geq 3F + 1$. Thus, fewer adversaries can be tolerated in the network compared to the malicious-type case with $n \geq 2F + 1$ given in the proposition above. This is intuitive since Byzantine adversaries are more powerful. For Proposition 1, we have proved using arguments similar to those in (LeBlanc et al. (2013); Dibaji and Ishii (2015)).

The advantage of our approach is that it can be extended to non-complete graphs as shown in the theorem below.

Theorem 1. Consider the multi-agent system under the network $\mathcal{G}$ where the mobile malicious agents follow the F-total and M1 model. Then, the regular agents using Protocol 1 reach resilient consensus if

C1 $n \geq 4F + 4$.
C2 For every agent i, the number of neighbors satisfies $|\mathcal{N}_i| \geq 2F + 1 + n/2$.

The safety interval is given by $\mathcal{S} = [\underline{x}(0), \overline{x}(0)]$.

Note that condition C1 in the theorem is necessary for condition C2 to hold. This can be shown by using complete graphs. For the case $n = 4F + 3$, each agent clearly has $4F+2$ neighbors. However, the condition C2 is $|\mathcal{N}_i| \geq 2F + 1 + \lceil n/2 \rceil = 4F + 3$, and thus there is a contradiction. Further, when applied to complete graphs, this result also exhibits some conservatism. The bound in Theorem 1 is $n \geq 4F + 4$ whereas in Proposition 1, it is $n \geq 2F + 1$.

4. PROTOCOL 2 FOR THE M2 MODEL

We next show another protocol that is effective for the M2 model. This model is different from M1 in that the recovering agents do not send their values to neighbors since they are aware of having been infected. Hence, in the worst case under the M2 model, at each round, there can be F-total malicious agents and, in addition, F agents that do not send values.

Protocol 2. At each round k, regular agent $i \in \mathcal{R}(k)$ executes the following three steps:

1. *(Send)* If agent i is not recovering with the cured flag $\theta_i(k) = 0$, then it broadcasts its current value $x_i(k)$.
2. *(Collect)* It collect values $x_j(k)$ of neighbors $j \in \mathcal{N}_i$.
3. *(Update)* If the cured flag is $\theta_i(k) = 0$, then agent i sorts the received values and its own value in descending order. Otherwise (i.e., $\theta_i(k) = 1$), agent i is recovering and sorts only the received values. Agent i then deletes the F largest and the F smallest values. Finally, agent i updates its value by (1).

Similar to Proposition 1, we have the following result for networks in the complete graph forms.

Proposition 2. Consider the multi-agent system whose network $\mathcal{G}$ forms a complete graph. Suppose that the mobile malicious agents follow the F-total and M2 model. Then, the regular agents using Protocol 1 reach resilient consensus if and only if $n \geq 3F + 1$. The safety interval is given by $\mathcal{S} = [\underline{x}(0), \overline{x}(0)]$.

In the M2 model, there may be up to F cured agents (with $\theta_i(k) = 1$) that are not allowed to send their values to neighbors. Hence, each regular node may not receive data from some neighbors. Among the data received, up to F of them may be faulty. Protocol 2 is effective for this model since each regular agent deletes $2F$ neighbor values in Step 3. In comparison with M1, to guarantee its resilience for M2, F more neighbors for each agent are needed. This argument also holds for the result extended to non-complete graphs as shown in the following.

Theorem 2. Consider the multi-agent system under the network $\mathcal{G}$ where the mobile malicious agents follow the F-total and M2 model. Then, regular agents using Protocol 2 reach resilient consensus if

C1 $n \geq 6F + 4$.
C2 For every agent i, the number of neighbors satisfies $|\mathcal{N}_i| \geq 3F + 1 + n/2$.

The safety interval is given by $\mathcal{S} = [\underline{x}(0), \overline{x}(0)]$.

We discuss the differences between M1 and M2. Generally, the graph condition for M2 is stricter than that for M1 because the agents in the cured status complicate the system behavior. Moreover, adversary agents in M2 are

Table 1. Properties of mobile adversary models

Model	Timing to move	Aware when cured	Complete graph	Non-complete graph
M1	Send	–	$n \geq 2F+1$	$n \geq 4F+4$
M2	Any	Yes	$n \geq 3F+1$	$n \geq 6F+4$
M3	Any	–	$n \geq 4F+1$	$n \geq 8F+4$

more powerful as they can move at any step during the update rounds while in M1, they switch only at send steps.

The main feature of M2 is that once an adversary agent moves away, the recovering agent soon knows that it was infected and then avoids sending its value to neighbors. In practice, this feature may not be easy to attain as it requires the implementation of fault detection. To deal with such an issue, we discuss yet another mobile adversary model M3 in the next section. In this case, detection of cured agents is not needed. We propose another protocol to solve the resilient consensus problem for M3.

5. PROTOCOL 3 FOR THE M3 MODEL

We outline the resilient protocol for the M3 model. Mobile adversaries under this model are more powerful since the recovering agents do not know about their infection. Hence, they send their values during the cured round though they can be corrupted. In this respect, the recovering agents can be considered as additional F-total malicious agents in the network. Thus, at each round, the regular agents may receive at most $2F$ corrupted values.

Protocol 3 copes with additional malicious data from the M3 model. It is slightly modified from Protocol 1. Specifically, in Step 3 at each round, $4F$ values ($2F$ largest and $2F$ smallest) are removed while in Protocol 1, this number is $2F$. We will show that this protocol is effective to deal with the mobile malicious agents under M3.

Since more neighbors are deleted in Protocol 3, each regular agent needs more neighbors compared with networks for Protocols 1 and 2. By an analysis similar to those in previous sections, we have the following results. The first is concerned with networks in the complete graph form.

Proposition 3. Consider the multi-agent system whose network forms a complete graph. Suppose that the mobile malicious agents follow the F-total and M3 model. Then, the regular agents using Protocol 2 reach resilient consensus if and only if $n \geq 4F+1$. The safety interval is given by $\mathcal{S} = [\underline{x}(0), \overline{x}(0)]$.

We can further deal with the non-complete graph case as shown in the theorem below.

Theorem 3. Consider the multi-agent system under the network $\mathcal{G}$ where the mobile malicious agents follow the F-total and M3 model. Then, regular agents using Protocol 3 reach resilient consensus if

C1 $n \geq 8F+4$.
C2 For every agent i, the number of neighbors satisfies $|\mathcal{N}_i| \geq 4F+1+n/2$.

The safety interval is given by $\mathcal{S} = [\underline{x}(0), \overline{x}(0)]$.

We highlight the differences of the M3 model and the related results from those for M1 and M2. First, we discuss the relation to M1. Both M1 and M3 do not require the functionality to detect agents in cured status. However, the M3 model is more powerful since in M1, the adversary agents can move only at the send step, while in M3, they can move at any step. This difference results in a more restrictive condition on the network structure to guarantee resilient consensus. We observe that each agent needs $2F$ more neighbors in M3 compared with M1. Next, we compare the M3 model with M2. In both M2 and M3, the adversary agents can move at any step. The difference comes from the detection ability in the regular agents, and the agents in M2 are more capable in this respect. In M2, if a regular agent is infected by an adversary, it becomes aware as soon as the adversary moves away. In contrast, the regular agents in M3 will never be aware of the infection. As discussed above, we can find that the graph conditions are related to the power of adversaries and defenders. Table 1 presents the models' properties.

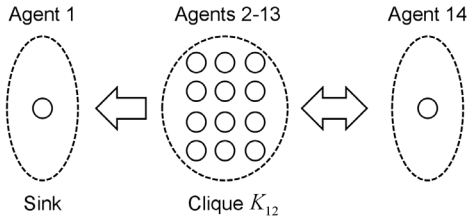


Fig. 2. Network structure in the simulations

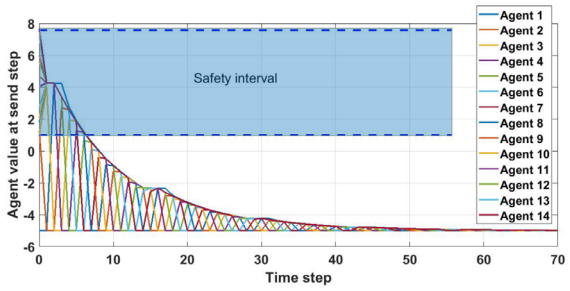
6. ILLUSTRATIVE EXAMPLE

We consider the network of 14 agents in Fig. 2. The subgraph consisting of agents 2 to 13 form a clique (a complete subgraph). All of them have edges to agent 1, whereas agent 1 has no outgoing edge to any agents and is a sink. Agent 14 is an incoming/outgoing neighbor of all nodes in the clique of agents 2 to 13, but not with agent 1.

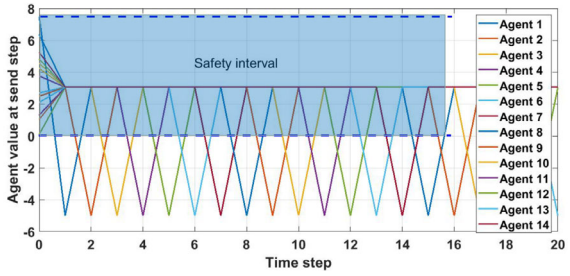
All agents have 12 neighbors, implying that resilient consensus can be attained under the three mobile adversary models. For the M1 model, Protocol 1 can tolerate up to two mobile malicious agents by Theorem 1. On the other hand, in the M2 and M3 models, Protocols 2 and 3 can handle one mobile malicious agent, respectively, by Theorems 2 and 3. Note that under the static model, this network may have up to $F = 4$ malicious agents because the network is $(5,5)$-robust. This difference indicates the difficulty to deal with mobile malicious agents.

In all simulations, we introduce one mobile malicious agent in the network and set $F = 1$. The agents' initial states are taken randomly from the interval $[0, 8]$. Thus, the safety interval $\mathcal{S}$ should be nonnegative. The mobile adversary agent moves from agent 1 to agent 14 in a round robin fashion. At each agent, during the infection, the state will switch to the same negative number -5.

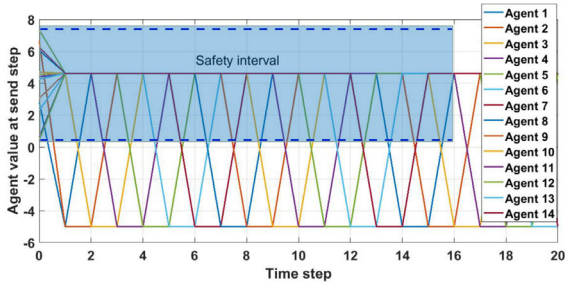
We first demonstrate the limitations of the conventional MSR algorithms from, e.g., (LeBlanc et al. (2013); Dibaji and Ishii (2015)) under the mobile adversary models. Here, the case for the M1 model is examined. The time responses of the agents' values are shown in Fig. 3(a). Though all regular agents eventually reach consensus, the value is -5. Hence, the mobile adversary agent is successful in leading all regular agents to a value that it specified. For the other two mobile adversary models, we obtain similar results.



(a) Conventional MSR algorithm under M1



(b) Protocol 1 under M1



(c) Protocol 3 under M3

Fig. 3. Time responses of agent values

Next, we check the effectiveness of the proposed protocols. We first examine Protocol 1 under the M1 model. The time responses are depicted in Fig. 3(b). Observe that each time the adversary agent moves away, the agent can recover in the next round and go back to take a state value as other regular agents. For Protocol 2 under the M2 model, we can obtain similar responses (not shown here).

For Protocol 3 under the M3 model, each regular agent removes more neighbor values at updates, so that the detection of cured agents is not needed. The results are displayed in Fig. 3(c), where resilient consensus is clearly reached. Notice that this plot is different from the one in Fig. 3(b) in that the values of each agent that turned malicious are kept at -5 for two time steps in a row. As a result, there are two agents taking false values at a time.

7. CONCLUSION

In this paper, we have considered resilient protocols for the multi-agent consensus problem to mitigate the influence of mobile misbehaving agents. By restricting mobile adversaries to malicious types, we have proposed three novel protocols for three mobile adversary models. We have derived conditions on the network structures for achieving resilient consensus for both complete and non-complete graphs. Future works will focus on formulating a more detailed model for mobile adversary behaviors.

REFERENCES

Bonnet, F., Défago, X., Nguyen, T. D., and Potop-Butucaru, M. (2016). Tight bound on mobile Byzantine agreement. *Theoretical Computer Science*, 609, 361–373.

Bonomi, S., Pozzo, A. D., Potop-Butucaru, M., and Tixeuil, S. (2019). Approximate agreement under mobile Byzantine faults. *Theoretical Computer Science*, 758, 17–29.

Bouzid, Z., Potop-Butucaru, M., and Tixeuil, S. (2010). Optimal Byzantine-resilient convergence in unidimensional robot networks. *Theoretical Computer Science*, 411, 3154–3168.

Buhrman, S., Garay, J. A., and Hoepman, J. H. (1995). Optimal resiliency against mobile faults. In *Proc. 25th Int. Symp. Fault-Tolerant Computing*, 83–88.

Chen, Y., Kar, S., and Moura, J. M. F. (2018). Resilient distributed estimation: Sensor attacks. *arXiv:1709.06156v2*.

Dibaji, S. M. and Ishii, H. (2015). Consensus of second-order multi-agent systems in the presence of locally bounded faults. *Systems & Control Letters*, 79, 23–29.

Dibaji, S. M., Ishii, H., and Tempo, R. (2018). Resilient randomized quantized consensus. *IEEE Trans. Automatic Control*, 63(8), 2508–2522.

Garay, J. A. (1994). Reaching (and maintaining) agreement in the presence of mobile faults. In *Proc. 8th Int. Workshop on Distributed Algorithms*.

Kieckhafer, R. M. and Azadmanesh, M. H. (1994). Reaching approximate agreement with mixed-mode faults. *IEEE Trans. Parallel Distributed Systems*, 5(1), 53–63.

LeBlanc, H. J., Zhang, H., Koutsoukos, X., and Sundaram, S. (2013). Resilient asymptotic consensus in robust networks. *IEEE J. Selected Areas Comm.*, 31, 766–781.

Lynch, N. A. (1996) *Distributed Algorithms*. Morgan Kaufmann.

Mesbahi, M. and Egerstedt, M. (2010). *Graph Theoretical Methods in Multiagent Networks*. Princeton Univ. Press.

Nowzari, C., Preciado, V. M., and Pappas, G. J. (2016). Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems Magazine*, 36(1), 26–46.

Sandberg, H., Amin, S., and Johansson, K. H. (Guest Eds) (2015). Special issue on cyberphysical security in networked control systems. *IEEE Control Systems Magazine*, 35(1), 2015.

Sasaki, T., Yamauchi, Y., Kijima, S., and Yamashita, M. (2013). Mobile Byzantine agreement on arbitrary network. In *Proc. 17th Int. Conf. Principles of Distributed Systems*, 236–250.

Vaidya, N. H., Tseng, L., and Liang, G. (2012). Iterative approximate Byzantine consensus in arbitrary directed graphs. In *Proc. ACM Symp. Principles of Distributed Computing*, pp. 365–374, 2012.

Wang, Y. and Ishii, H. (2019). An event-triggered approach to quantized resilient consensus. In *Proc. European Control Conference*, 2719–2724.

Wang, Y. and Ishii, H. (2020). Resilient consensus through event-based communication. *IEEE Trans. Control of Network Systems*, to appear.

Zhang, H., Fata, E., and Sundaram, S. (2015). A notion of robustness in complex networks. *IEEE Trans. Control of Network Systems*, 2(3), 310–320.