2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering

# Resilient Bipartite Consensus for Multi-Agent Networks with Antagonistic Interactions

Hongbin Liu*, Ming Xu*†, Yiming Wu✉†, Ning Zheng*, Yuanfang Chen†, Md Zakirul Alam Bhuiyan‡
*School of Computer Science and Technology, Hangzhou Dianzi University, China
†School of Cyberspace, Hangzhou Dianzi University, China
‡Department of Computer and Information Sciences, Fordham University, USA

***Abstract*— This paper addresses the problem of resilient bipartite consensus for multi-agent networks in the presence of misbehaving nodes. The bipartite consensus problem was first studied by C. Altafini, and has been extensively studied in recent years. The interaction representing the communication between two agents is characterized by edge weights in a signed directed graph where the positive weight of an edge implies cooperation between the two agents while a negative one corresponds to antagonism. Resilient consensus problems without antagonistic interactions have been exhaustively studied, while the security problem of bipartite consensus, to the best of our knowledge, has not been studied yet. In this paper, we extend the resilient consensus problem to the case when antagonistic interactions exist. The developed results which are applicable for multi-agent systems with continuous-time dynamics shows that all normal nodes reach resilient bipartite consensus if the associated signed digraph is structurally balanced and has sufficient connectivity in terms of robustness. Numerical examples are provided to illustrate our results.**

***Index Terms*—Bipartite consensus, signed graph, resilient consensus, antagonistic interactions, malicious attack.**

## I. Introduction

In recent years, the consensus problems of multi-agent systems are of great academic vitality, mainly owning to its wide applications in various areas such as distributed classification, distributed optimization, formation control, and sensor networks. While things become more complicated in consensus problems because of the malicious attackers in recent years. Compromised nodes in the system may cause irreparable harm to the whole consensus process. Therefore, it is important to implement embedded security consensus algorithms in multi-agent systems to make safe and reliable performance possible.

Reaching consensus in the presence of faulty or misbehaving nodes has received significant attention from the research community [1]–[6]. In some researches such as [3], [7], connectivity, which is one of the conventional graph theoretic property is utilized to study the secure robustness of a certain network. Their results show that a network with at most $F$ malicious nodes can reach consensus if the connectivity of the network is no less than $2F+1$. However, these consensus algorithm usually do not suits large-scale networks since they either require that the normal nodes have more or less some non-local information, or need the topology of the network to be complete, i.e., a complete graph is needed, and both of them cause more resource consumption. Local communication is more suitable for a large-scaled network since it requires less communication between agents. In [8], the authors first propose an algorithm using only local information to solve the consensus problem in the presence of Byzantine faults in finite time. To enhance resiliency, LeBlanc *et al.* [4] propose a novel topological property, named *network robustness*, and provide a comprehensive characterization of the network topology. The network robustness suits the algorithm that only local information is needed, such as weighted mean subsequence reduced (W-MSR) algorithms [9], [10]. The main idea of W-MSR algorithms is that each normal node removes the extreme values with respect to its own value. Then, combining the ideas from the robustness property and W-MSR algorithms, [4] propose a resilient consensus strategy for *Byzantine* nodes under both $F$-local and $F$-local thread models. While W-MSR used in [4] may not suitable when antagonistic interactions exist, so we extend it to AW-MSR algorithm(absolute weighted mean subsequence reduced, see Section V-A for details) to suit our case.

The previous consensus algorithms which are originated from iterative procedures of decision-making, are based on a cooperation interaction between agents [11]–[13], i.e., all agents consider its neighbors are “friends”. While, motivated by opinion dynamics over social networks [14]–[16], this “cooperation” idea of consensus has been extended to a more general multi-agent systems which allows both cooperation and antagonistic relationships. It is common in many antagonistic systems such as two-party political systems, rival business cartels, and teams opposed in a sport match, etc. Altafini start the research on consensus algorithm for multi-agent networks with antagonistic interactions [17]. In his work, the network modeled by a signed digraph, and a specific consensus which establishes *bipartite consensus* is reached, where all values of the agents are the same in modulus but are different in sign. However, the security problem of bipartite consensus is not considered by Altafini, he only deals with the bipartite consensus issues. The security problem of bipartite consensus has not attracted sufficient attention.

In this paper, we extend the resilient consensus problem to the case where antagonistic interactions exist, or we can say

Corresponding author: Y. Wu (*E-mail: yimgwu@hotmail.com*).

2324-9013/18/31.00 ©2018 IEEE
DOI 10.1109/TrustCom/BigDataSE.2018.00174

that we extend the bipartite consensus problems to the case when attack exists. Our aim in this paper is to characterize the structure of the network topology necessary and sufficient to achieve bipartite consensus in the presence of malicious nodes. We prove that it can be transfered to a standard consensus problem so that we can solve it by using existing methods and theories, and we show that given a structurally balanced digraph, resilient bipartite consensus can be reached under $F$-local attack model if it has sufficient network robustness.

This paper is organized as follows. After discussing related work in Section II, we introduces the relevant notations and preliminaries which are needed in the paper in Section III. In Section IV, we describe the problem formulation. Resilient consensus algorithm is introduced in section V. Section VI presents the main results. Some numerical example are presented for illustration in Section VII. Finally, Section VIII concludes the paper.

## II. Related Work

Resilient consensus is a special case of distributed consensus. It is assumed that a limited number of malicious nodes are present that aim to disrupt consensus process by providing false data to neighboring nodes [18]. The existing literature most closely related to this paper are the first order resilient consensus results of [4], [19], [20], and the second order resilient consensus results of [21], [22]. In [4], [5], [22], resilient consensus problems are considered under a discrete-time model. While in [20], [23], [24], resilient consensus problems are considered under a continuous-time model. [23] and [25] have studied resilient consensus problem with quantized communication network. Recently, the notion of $r$-robustness in graphs, introduced in [19] and [20] has received much attention in characterizing resilience of consensus process for multi-agent networks in adversarial environments. In [10], the authors have shown that $r$-robustness property of the graph in general is a stronger certificate of structural robustness as compared to the network connectivity. Contrary to the approach of achieving desired $r$-robustness in graphs by strategically adding edges [26]. The author in [27] use the notion of trusted nodes, and show that by selecting a small subset of nodes as trusted, one can achieve any desired value of $r$-robustness.

Different from above literature that all the agents achieve their common goals due to cooperation, antagonism is also common in real-world networks. Altafini [17] first study the consensus problem with antagonistic interactions. He show that on a signed network, all nodes could reach bipartite consensus if the effective connectivity condition is satisfied. Since then, the bipartite consensus in a finite time and the bipartite consensus under switching topology network are studied by [28] and [29] respectively.

## III. Notations And Preliminaries

In this section, we introduce some relevant notations and preliminaries that are needed for signed digraphs and bipartite consensus problem. In addition, we introduce some concepts of network robustness which is useful for our problems.

### A. Notations

Throughout this paper, we denote the set of integers by $\mathbb{Z}$, and the set of real numbers by $\mathbb{R}$. $m : n (m, n \in \mathbb{Z})$ stands for the index set $\{m, m+1, ..., n\}$, where $m \leq n$. $1_n = [1, 1, ..., 1]^T \in \mathbb{R}^n$, and $diag(d_1, d_2, ...d_n)$ stands for a diagonal matrix whose diagonal entries are $d_1, d_2, ...d_n$, and off-diagonal entries are all zero. We divide all the labels denoted by $P$ of nodes of the network into two parts, $I$ stands for the *normal* part, and $M$ stands for the *malicious* part. Obviously, $I \cup M = P, I \cap M = \emptyset$. Let $\mathcal{D} = \{D = diag(\zeta), \zeta = [\zeta_1, ...\zeta_n], \zeta_i \in \{1, -1\}\}$ be the set of all diagonal matrix whose diagonal entries are limited to {1, -1}. We denote $sign(x)$ as the sign function of a scalar x $\in \mathbb{R}$, i.e.,

$$sign(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases} \quad (1)$$

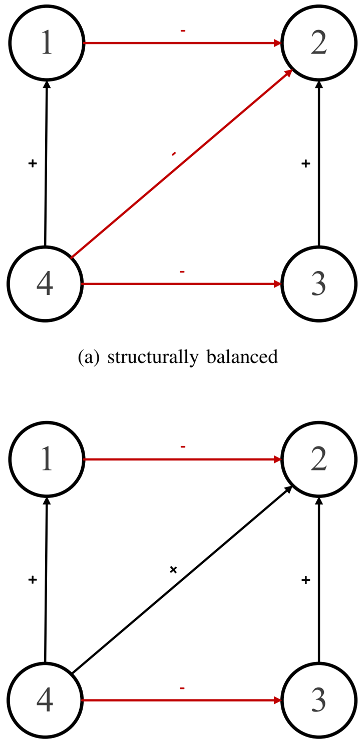
Additionally, we denote $G(A)$ as a signed digraph whose adjacency weight matrix is $A$.

### B. Signed digraphs

The interaction network can be modeled by a signed digraph(short for "directed graph") which is represented by a triple $G = (V, E, A)$, consisting of a vertex set $V = \{v_1, v_2, ..., v_n\}$, an edge set $E \subseteq V \times V = \{(v_i, v_j) : v_i, v_j \in V\}$ which is defined such that $(v_j, v_i)$ is a directed edge from $v_j$ to $v_i$, i.e., vertex $v_j$ is a neighbor of vertex $v_i$, and an adjacency weight matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ which is defined such that $(v_j, v_i) \in E \Leftrightarrow a_{ij} \neq 0$, and otherwise, $a_{ij} = 0$. Let $N_i = \{j : (v_j, v_i) \in E\}$ denote the set of labels of those vertices that are neighbors of $v_i$. A digon in a digraph is a pair of edges sharing the same nodes $(v_i, v_j), (v_j, v_i) \in E$. We call a graph is digon sign-symmetric [17] if $a_{ij}a_{ji} \geq 0, \forall i, j \in 1 : n$, which means that the edge pairs of all digons cannot have opposite signs. We assume that throughout this paper, graph $G$ has no self-loops, i.e., $a_{ii} = 0, \forall i \in 1 : n$, and $G$ is digon sign-symmetric.

*Definition 1:* Given a signed digraph $G$, we say $G$ is structurally balanced if there is a partition $\{V_1, V_2\}$ of its set of vertices $V$, $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$, where all edges within $V_1$ and $V_2$ are positive while all edges between $V_1$ and $V_2$ are negative. Formally, $a_{ij} \geq 0$ for $\forall v_i, v_j \in V_p (p \in \{1, 2\})$ and $a_{ij} \leq 0$ for $\forall v_i \in V_p, v_j \in V_q, p \neq q (p, q \in \{1, 2\})$, otherwise, $G$ is structurally unbalanced.

A structurally balanced graph implies that a community is divided into two hostile camps, such as votaries of two political parties, or teams opposed in a sport match. They cooperates with members within his own camp while competes with his opponents. The following lemma gives the sufficient conditions of signed graph to be structurally balanced:



(a) structurally balanced

(b) structurally unbalanced

Fig. 1: Illustration of structurally balanced and unbalanced graphs. Fig. 1(a) can be divided into two blocks $V_1 = [1, 4]$ and $V_2 = [2, 3]$ and the edges inside $V_1$ and $V_2$ are positive, while edges between $V_1$ and $V_2$ are negative. While Fig. 1(b) cannot since edge between node 4 and 2 is positive.

*Lemma 1:* We say a signed digraph $G$ is structurally balanced if there exists a matrix $D \in \mathcal{D}$ such that $D^{-1}AD$ is nonnegative.

*Proof:* According to definition 1, we can easily choose a $D = diag(\zeta), \zeta = [\zeta_1, ...\zeta_n], \zeta_i \in \{1, -1\}$ such that $\zeta_i = +1$ when $v_i \in V_1$ and $\zeta_i = -1$ when $v_i \in V_2$ to obtain a transformed matrix $DAD$ that has all nonnegative entries.

As depicted in Fig. 1, a structurally balanced signed digraph can be divided into two blocks $\{V_1, V_2\}$, here $V_1 = [1, 2], V_2 = [3, 4]$, all positive edges connect nodes within $V_1$ or $V_2$, and negative edges connect nodes between $V_1$ and $V_2$. Here Fig. 1(a) is structurally balanced while Fig. 1(b) is structurally unbalanced.

Obviously, we can notice that there is a node remove operation in AW-MSR(see V-A for details) algorithm which means that the interaction network is changing over time. So we need the following definition:

*Definition 2:* The joint graph of $G$ during time interval $[t_1, t_2]$ is defined by $G[t_1, t_2] = \bigcup_{t\in[t_1,t_2]} G_t = (V, \bigcup_{t\in[t_1,t_2]} E_t)$. $G[t_1, t_2]$ is called uniformly jointly quasi-strongly connected(UQSC) if there exists a constant $T \geq 1$ such that $G[t, t + T]$ has a spanning tree for any $t \geq 0$, and $G[t_1, t_2]$ is sign consistent.

A sign consistent joint graph $G[t_1, t_2]$ means that the nodes in the two hostile camps will not change during interval $[t_1, t_2]$ given that the signed joint graph is structurally balanced.

### C. Bipartite Consensus

In this paper, we consider a interaction multi-agent network modeled by a signed digraph $G$ which consists of $n$ agents indexed 1 through $n$. Each agent is regarded as a vertex in $G$, and the opinion of the $i$-th agent is denoted by $x_i \in \mathbb{R}$, so the opinion of all the agents can be defined as a vector $x = [x_1, x_2, ...x_n]^T \in \mathbb{R}^n$. We say a multi-agent system reaches bipartite consensus if all its $n$ agents converge to values that are the same in modulus but different in sign. More specifically, we have the following definition:

*Definition 3:* The system reaches a bipartite consensus, if for any $x(0)$ (initial state of all agents), there exists $x^* > 0$ such that

$$\lim_{t\to+\infty} |x_i(t)| = x^*, i \in 1 : n, x^* \in \mathbb{R}, \tag{2}$$

where $x_i(t) \in \mathbb{R}$ is the state of agents at time $t$, and $x^* \in \mathbb{R}$ refers to their final state.

### D. Network Robustness

Traditional secure and fault-tolerant consensus algorithms typically assume knowledge of nonlocal information which is not suitable for large-scale dynamic networks. In [4], the authors proposed a novel graph-theoretic property which called network robustness, and designed consensus algorithm using only local information that is resilient to faults and compromised nodes. This robustness notion is very useful in characterizing resilience of various dynamical processes over networks in adversarial environments.

*Definition 4:* ($r$-reachable set [4]): Given a digraph $G = \{V, E\}$ and a nonempty subset $S \subset V$, we say $S$ is a $r$-reachable set if there exists at least one node $i \in S$ that at least $r$ nodes in $N_i$ comes from outside $S$, i.e., $|N_i \setminus S| \geq r$, where $r \in \mathbb{Z} \geq 0$.

As illustrated in Fig. 2, set $S$ is $r$-reachable if it contains at least one node that has at least $r$ neighbors outside of $S$. It means that at least one node inside $S$ can obtain information from a certain number of nodes out of $S$ and parameter $r$ quantifies the information flow. The $r$-reachability property pertains to a given set $S$, and we also need a generalization definition of this notion of redundancy for the entire interaction network.

*Definition 5:* ($r$-robustness [4]): Given a digraph $G$, we say $G$ is $r$-robust if for every pair of nonempty, disjoint subsets of $V$, denoted by $S_1, S_2$, at least one of them is $r$-reachable.

By employing the notion of robustness, some properties of the $r$-robust graph are recalled below.

*Lemma 2:* Consider an $r$-robust graph $G = (V, E)$. Let $\hat{G}$ be the graph generated by removing up to $s(s < r)$ incoming edges of each node of $V$, then, we say that $\hat{G}$ is an $(r - s)$-robust graph.

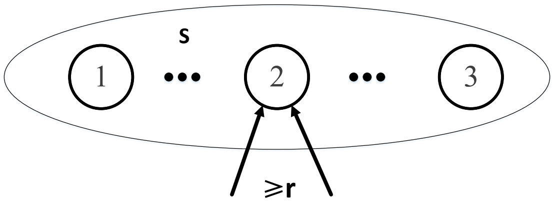


Fig. 2: Illustration of an $r$-reachable set of nodes. Set $S$ is $r$-reachable since node 2 has more than $r$ neighbors from outside $S$.

*Lemma 3:* Graph $G$ contains a spanning tree if and only if $G$ is 1-robust.

*Proof:* sufficiency: By contradiction, we assume that $G$ contains a spanning tree while it is not 1-robust. According to definition 5, their exists two disjoint subsets, $S_1$ and $S_2$, who do not have neighbors outside their own set. Then their is no information flow between $S_1$ and $S_2$, which contradicts the definition of spanning tree.

necessity: Similarly, we assume that $G$ does not contain a spanning tree. The adjacency matrix $A$ of $G$ is decomposable according to [30], [31], which means their exists two sets $S_1 \cup S_2 = V$ and $S_1 \cap S_2 = \emptyset$, and their are no information flow between $S_1$ and $S_2$ which contradicts the definition of 1-robust.

### E. Laplacian matrix

To establish our results, we need to introduce some properties of laplacian matrix. Given a signed digraph $G$, and its adjacency matrix $A$, the following lemma disclose the relationship between the eigenvalue distribution of the laplacian matrix of $A$ and the spanning tree property of $G$.

*Lemma 4:* If $G$ is structurally balanced, then 0 is the simple eigenvalue of $L$ and all the other eigenvalues have positive real parts if and only if $G$ contains a spanning tree.

*Proof:* Following from Gershgorin's disc theorem [32], for a directed graph $G$, all of the nonzero eigenvalues of $L$ of $G$ have positive real parts. The proof follows immediately from the following equivalent conditions:

1) Graph $G$ contains a spanning tree and structurally balanced.
2) $\exists D \in \mathcal{D}$ such that $D^{-1}AD$ has all nonnegative entries and $G$ contains a spanning tree.
3) $\exists D \in \mathcal{D}$ such that $D^{-1}LD$ has a simple eigenvalue 0 and its corresponding eigenvector is $\mathbf{1}$.

Condition 1) and condition 2) are equivalent according to Lemma 1. Note that laplacian matrix $L$ has zero row sums, 0 is thus an eigenvalue of $L$ with its associated eigenvector $\mathbf{1}$, where $\mathbf{1} = [1, ..., 1]^T$. Condition 3) and condition 2) are equivalent since $L$ and $D^{-1}LD$ are isospectral, i.e., $sp(L) = sp(D^{-1}LD)$, according to proposition 2 of [17].

## IV. Problem Formulation

In this section, we formulate our problems. Firstly, we introduce the update model, i.e., the consensus protocol in the presence of antagonistic interactions. This is different from conventional consensus protocols which only consider the case of cooperation interactions. And therefore, we need a new definition of laplacian matrix corresponds to the signed network. Secondly, the thread model definition is given in the next subsection. Finally, we formally define the concept of resilient bipartite consensus.

### A. Update Model

As mentioned above in section III, an adjacency matrix $A \in \mathbb{R}^{n\times n}$ can be assigned to a signed weighted digraph which can represented by $G(A)$. The Laplacian matrix is defined as follow:

$$L = C - A$$

where

$$C = diag(c_1, ..., c_n), c_i = \sum_{j=1}^{n} |a_{ij}|, i \in 1:n,$$

and $a_{ij}$ are assumed to satisfy the following assumption.

*Assumption 1:* For each $i \in 1:n$, there hold $a_{ii} = 0$ and

$$\sum_{j=1}^{n} |a_{ij}| = 1,$$

for all time $t$. There exists a number $\alpha > 0$ such that $|a_{ij}(t)| \geq \alpha$ when $|a_{ij}(t)| > 0$ for all $i, j \in 1:n$ and $t$.

The elements in $L$ is therefore,

$$l_{ij} = \begin{cases} \sum_{j\in N_i} |a_{ij}|, & k = i \\ -a_{ik}, & k \neq i. \end{cases} \tag{3}$$

The normal agents update their opinions in accordance with a distributed protocol as follows:

$$\dot{x}(t) = -L[A(t)]x(t), t \geq 0, \tag{4}$$

which in components reads

$$\dot{x}_i(t) = \sum_{k\in N_i} |a_{ik}(t)|(sign(a_{ik}(t))x_k(t) - x_i(t)), i \in I, \tag{5}$$

while the compromised nodes update their opinions in a arbitrary function $f_i'(*), i \in M$.

### B. Thread Model

Usually, the thread model in multi-agent systems consist of two aspects, one is the method of attack and the other one is the scope of thread.

*Definition 6:* A node $v_i \in V$ is said to be *Byzantine* if it has the following features:

1) sends different values to different neighbors at the same time.
2) updates its value in an arbitrary function, i.e., $f_i'(*), i \in M$.
3) change its attack target or abandon at any time.

*Definition 7:* A node $v_i \in V$ is said to be *malicious* if it satisfy the following features:

1) sends the same values to all its neighbors at the same time.

2) updates its value in an arbitrary function, i.e., $f'_i(*), i \in M$.

Note that Byzantine nodes are more harmful compared to malicious nodes because they cover almost all the features that other kinds of attacks possess. So we say an algorithm is resilient to almost all other attack if it is resilient to Byzantine attack. Having defined the type of misbehavior in the network, it is necessary to define the $number$ of misbehaving nodes, i.e., the scope of threads.

*Definition 8:* ($F$-total set [4]): A set $S \subset V$ is said $F$-total if it contains at most $F$ nodes in the whole network, i.e., $|S| \leq F$ where $F \in \mathbb{Z} \geq 0$.

*Definition 9:* ($F$-local set [4]): A set $S \subset V$ is said $F$-local if it contains at most $F$ nodes in the neighborhood of the other nodes for all $t$, i.e., $|N_i[t] \cap S| \leq F, \forall i \in V \setminus S, \forall t \in \mathbb{Z} \geq 0$.

*Definition 10:* A set of adversary nodes is $F$-totally bounded or $F$-locally bounded if it is an $F$-total set or $F$-local set respectively. We refer to these thread scopes as the $F$-total or $F$-local models, respectively.

Obviously, if there is no number limitation of the misbehaving nodes, it's hard for the multi-agent system to achieve consensus since there may be too many adversary nodes. And it is also difficult for us to analyze the security issues.

### C. Resilient Bipartite Consensus

Given the thread model and scope of threads, we now formally define resilient bipartite consensus.

*Definition 11:* (Resilient Bipartite Consensus): The system reaches a resilient bipartite consensus under $F$-total model or $F$-local model attack if all the normal nodes reaches a bipartite consensus(see III-C), i.e,

$$\lim_{t \to +\infty} |x_i(t)| = x^* > 0, i \in I, x^* \in \mathbb{R}, \tag{6}$$

for any choice of initial values.

Different from resilient asymptotic consensus in [4], here resilient bipartite consensus requires an *absolute* agreement, i.e., they converge to values that the same in modulus but different in sign.

## V. Resilient Consensus Algorithm

Their are various approaches to facilitate consensus in conventional consensus problems under attacks. Here we extend W-MSR algorithm to AW-MSR algorithm so that it can suit the case when antagonistic interactions exist. Character A stands for "absolute", means that the weight in new version update rule (see step 3)) is different, we need its absolute value.

### A. AW-MSR Algorithm

In fact, AW-MSR only makes some modification on the basic update rule. At every time-step $t$, each normal node $i$ obtains the values from its neighbors. Under the definition of thread model and scope of thread, at most $F$ nodes of $i$'s neighbors may be misbehaving. Unfortunately, node $i$ cannot determine which neighbor(s) may be compromised. The core idea of AW-MSR is that each node removes the extreme values received from its neighbors with respect to its own value at every time-step, while when the value of malicious node is **not** out of a certain bounds(with respect to all $i$'s neighbors and its own value), and it may not be removed, it does not affect very much since it misbehaves just like a normal node at current time-step. Specifically,

1) Each node obtains values from its neighbors according to update rule(equation 5), and forms a sorted value list $values$.
2) Remove the nodes whose values are strictly larger or smaller than its own value $x_i(t)$, but only $F$ can be removed respectively. More specifically, if there are less than $F$ values strictly larger(or smaller) than its own value, then, remove all values that are strictly larger(or smaller) than its own. Otherwise, remove precisely the largest(or smallest) $F$ values in $values$.
3) Each normal node updates according to the new version of rule

$$\dot{x}_i(t) = \sum_{k \in N_i \setminus R_i(t)} |a_{ik}(t)| \times (sign(a_{ik}(t))x_k(t) - x_i(t)), i \in I. \tag{7}$$

Here $R_i(t)$ stands for the set of nodes removed by node $i$ in at time-step $t$ in step 2).

Note that the weight of an edge is now different from that when only cooperation exists. Here we assume that $a_{ik}(t)$ satisfies assumption 1 and

$$|a_{ik}(t)| = 1/|N_i \setminus R_i(t)|,$$

where $|N_i \setminus R_i(t)|$ is the number of set $N_i \setminus R_i(t)$. We consider $F$ as the parameter of AW-MSR under $F$-local or $F$-total model. Note that the set of nodes removed by node $i$, $R_i(t)$, is possibly time-varying. Hence, AW-MSR effectively induces switching behavior even though the topology of the interaction network is fixed. When in the case of a signed digraph, a topology switching may change the connectivity property of the network, and it may not satisfy the conditions to achieve the bipartite consensus unless it has sufficient connectivity in terms of robustness.

## VI. Main Results

This section starts with the case of signed network under no attack and the network is time-invariant, where sufficient and necessary conditions are given for reaching bipartite consensus. Note that the network that AW-MSR algorithm is not applied is time-invariant, which means that its corresponding adjacency matrix $A$ is static, i.e., $A(t) \equiv A(0)$. When AW-MSR algorithm is applied, the situation may be different since AW-MSR causes a topology switching, i.e., the network is time-varying. So, we also give the result of the case when the network is time-varying(AW-MSR is applied). We then make the network expose in a $F$-local attack, and see its converge process. Finally, we check the performance of AW-MSR with parameter $F$ under $F$-local attack. Note that here we only consider the $F$-local model since $F$-total model are similar to it.

*Theorem 1:* Protocol (4) establishes bipartite consensus if and only if $G[A]$ is structurally balanced and contains a spanning tree.

*Proof:* Given that graph $G$ is structurally balanced, there exists a diagonal matrix $D = diag(d_1, ..., d_n) \in \mathcal{D}$, where $d_i = 1$ if $i \in V_1$ and $d_i = -1$ if $i \in V_2$. Note that we only consider the case when $V_1 \neq \emptyset$ and $V_2 \neq \emptyset$. According to lemma 4, applying a *gauge transformation* $D$, we have

$$z(t) = Dx(t),$$

since $D^{-1} = D, x = D^{-1}z$, then, system (4) is transformed into

$$\dot{z}(t) = -L[|A|]z(t), t \geq 0, \tag{8}$$

where

$$L[|A|] = L_D = C - DAD,$$

$L_D$ is the new laplacian matrix corresponds to adjacency matrix of $G$. Note that this is a standard consensus problem, according to [33], we have

$$\lim_{t \to +\infty} z(t) = v^T z(0)\mathbf{1} = v^T D^{-1} x(0)\mathbf{1},$$

where $v = [v_1, v_2, ..., v_n]^T$ is the nonnegative left eigenvector of $D^{-1}LD$ and its corresponding eigenvalue is 0 [33], and $v^T\mathbf{1} = 1$, i.e., $v^T(D^{-1}LD) = 0$. So, the consensus solution of system (4) is

$$\lim_{t \to +\infty} x(t) = v^T D^{-1} x(0) D\mathbf{1},$$

Therefore protocol (8) establishes consensus, corresponding to bipartite consensus of protocol (4).

Given a set of structurally balanced signed digraphs with the same weights but different signs. For the adjacency matrix $A$ of each one of them, we can find a matrix $D \in \mathcal{D}$ such that $D^{-1}AD$ is nonnegative, means that all these networks are related by gauge transformations and all are isospectral, i.e., the corresponding Laplacians shares the same convergence processes. Thus applying a gauge transformation will not change their absolute values and convergence processes.

Theorem 1 considers the static case when their are no attack. While when AW-MSR algorithm is used, the graph topology is time-varying because of its node remove operation. Thus, we also give the following theorem of time-varying case.

*Theorem 2:* Protocol (4) establishes bipartite consensus if and only if $G[A(t)]$ is structurally balanced, sign consistent and UQSC.

*Proof:* Since the network is time-varying, we introduce a sequence of adjacency matrices $\{A_1, ..., A_m\}$ corresponding the graph at different time. Note that $G[A(t)]$ is sign consistent, so that if $A_i, i \in 1 : m$ is structurally balanced, $A_i, i \in 1 : m$ share the same matrix $D \in \mathbb{D}$ such that $DA_iD$ have all nonnegative entries. Similar to proof of theorem 1, Then we transfer protocol (4) into

$$\dot{z}(t) = -L[|A(t)|]z(t), t \geq 0, \tag{9}$$

where

$$L[|A(t)|] = L_{D(t)} = C - DA(t)D.$$

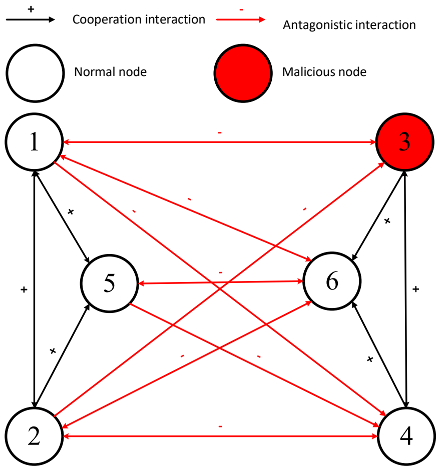


Fig. 3: A 3-robust graph (node 3 is malicious node). It can be divided into two parts $V_1 = [1, 2, 5]$ and $V_2 = [3, 4, 6]$. Nodes in $V_1$ and $V_2$ cooperate and nodes between $V_1$ and $V_2$ compete.

Note that this is also a standard consensus problem, and according to [33], we can obtain our results similar to theorem 1. Then we have the following corollary when AW-MSR algorithm is used.

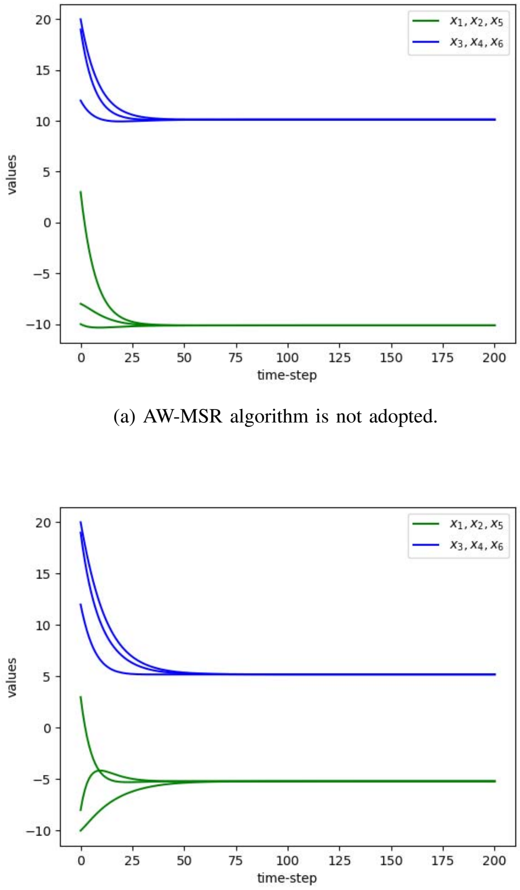
*Corollary 1:* Consider a time-invariant structurally balanced network modeled by a signed digraph $G = \{V, E, A\}$ where each normal node updates its value according to the AW-MSR algorithm with parameter $F$. Bipartite consensus can be achieved if the graph is $(2F$+1)-robust, and $V_1 = V_2 \neq \emptyset$.

*Proof:* According to lemma 2, the topology under AW-MSR with parameter $F$ is at least 1-robust, since each node removes at most $2F$ from its neighbors at time $t$. And $G[A(t)]$ contains a spanning tree since it is at least 1-robust according to lemma 3. Obviously, this is a time-varying case the same as theorem 2.

*Theorem 3:* Consider a time-invariant structurally balanced network modeled by a signed digraph $G = \{V, E, A\}$ where each normal node updates its value according to the AW-MSR algorithm with parameter $F$. Suppose assumption 1 holds, resilient bipartite consensus is achieved if the topology if the network is $(2F+1)$-robust under the $F$-local malicious model.

*Proof:* Similar to proof of theorem 1 and 2, we transfer system (4) to (8). By transforming the bipartite consensus problem into a standard consensus problem, we convert our problem into the case of a resilient consensus problem, which is just the same as the case in [4]. Note that the graph is $(2F$+1)-robust. Under assumption 1, we can conclude that graph $G$ can achieve bipartite consensus according to Theorem 2 of [4].

Note that our results can be easily extended to $F$-total attack model and $(r, s)$-robust cases in [4], here we omit for space limitation.



(a) AW-MSR algorithm is not adopted.

(b) AW-MSR algorithm is adopted.

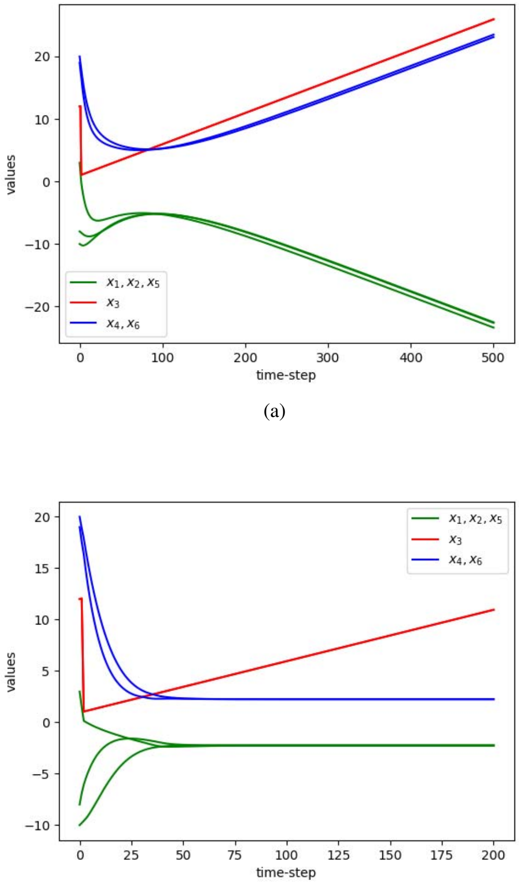
Fig. 4: Simulation results for the network without attack.

## VII. Simulations

In this section, we verify the theoretical results by using some simulations. We first show that the system cannot achieve bipartite consensus when AW-MSR algorithm is not used, while it establishes resilient bipartite consensus under AW-MSR with parameter $F$. We consider system (4) associated with a structurally balanced, 3-robust signed digraph depicted in Fig. 3, where the nodes are indexed 1 through 6. Suppose that node 3 is compromised and turns a malicious node. Its objective is to prevent the normal nodes from reaching bipartite consensus. It can be divided into two blocks: $V_1 = [1, 2, 5], V_2 = [3, 4, 6]$, in which nodes in $V_i, i \in \{1, 2\}$ cooperates and nodes in $V_i, V_j, i, j \in \{1, 2\}, i \neq j$ are opposed. It will cost a considerable amount of time to verify that Fig. 2 is 3-robust since there are no efficient algorithm so far to compute the robustness of a graph. We assume the initial state of this system is

$$x(0) = [-10, 3, 12, 19, -8, 20].$$

The results for the network of Fig. 3 are shown in Fig. 4 and Fig. 5. The green curves represents the consensus process of nodes in $V_1$ and blue curves for nodes in $V_2$, the red curve depicts the malicious(node 3) node's changing tendency over time. Fig. 4 shows the results of time-invariant and time-varying(under AW-MSR) cases. Theorem 3 implies that graph in Fig. 3 can sustain 1 malicious node under 1-local model. As previously supposed, node 3 is the malicious node. Obviously, node 3 is able to drive the normal nodes beyond its convergence process when AW-MSR is not applied while it fails whenever AW-MSR is used. Note that under AW-MSR, although resilient bipartite consensus can be reached, the consensus process are more or less affected by malicious nodes. The final state of all the nodes may be different when they are under no attack as we can see the difference between Fig. 4(a) and Fig. 5(b).



(a)

(b)

Fig. 5: Performance of the network under 1-local attack: (a) without AW-MSR algorithm and (b) with AW-MSR algorithm.

## VIII. Conclusion

The resilient consensus problem has attracted much attention due to its extensive applications in different areas. Various approaches have been proposed to facilitate security of consensus, while the case when antagonistic interactions exist has not been studied yet. In this paper, we have dealt with this problem by introducing the concept of network robustness and proved that the resilient bipartite consensus could be established given a structurally balanced and robust network. We showed that the resilient consensus problem in the presence of antagonistic interactions can be transfered into a standard consensus problem. The necessary and sufficient conditions for reaching bipartite consensus among non-faulty agents have been derived based on the graph topology notion.

## Acknowledgment

This work is supported by the cyberspace security Major Program in National Key Research and Development Plan of China under grant 2016YFB0800201, Natural Science Foundation of China under grants 61572165 and 61702150, State Key Program of Zhejiang Province Natural Science Foundation of China under grant LZ15F020003, Key Research and Development Plan Project of Zhejiang Province under grants 2017C01062 and 2017C01065, Zhejiang Provincial Public Research Project under grant LGG18F020015, and the Scientific Research Fund of Zhejiang Provincial Education Department under grant Y201737924.

## References

[1] F. Pasqualetti, A. Bicchi, and F. Bullo, "On the security of linear consensus networks," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC)*, 2009, pp. 4894–4901.

[2] ——, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 90–104, 2011.

[3] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1495–1508, 2011.

[4] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 4, pp. 766–781, 2013.

[5] Y. Wu, X. He, S. Liu, and L. Xie, "Consensus of discrete-time multi-agent systems with adversaries and time delays," *International Journal of General Systems*, vol. 43, no. 3-4, pp. 402–411, 2014.

[6] D. S M and H. Ishii, "Resilient consensus of second-order agent networks: Asynchronous update rules with delays," *IEEE Transactions on Automatic Control*, vol. 81, no. 2017, pp. 123–132, 2017.

[7] N. A. Lynch, *Distributed algorithms*. Elsevier, 1996.

[8] D. Dolev, N. A. Lynch, S. S. Pinter, E. W. Stark, and W. E. Weihl, "Reaching approximate agreement in the presence of faults," *Journal of the ACM (JACM)*, vol. 33, no. 3, pp. 499–516, 1986.

[9] M. M. De Azevedo and D. M. Blough, "Multistep interactive convergence: An efficient approach to the fault-tolerant clock synchronization of large multicomputers," *IEEE Transactions on Parallel & Distributed Systems*, vol. 9, no. 12, pp. 1195–1212, 1998.

[10] H. Zhang, E. Fata, and S. Sundaram, "A notion of robustness in complex networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 310–320, 2015.

[11] C.-L. Liu and F. Liu, "Dynamical consensus seeking of second-order multi-agent systems based on delayed state compensation ," *Systems & Control Letters*, vol. 61, no. 12, pp. 1235–1241, 2012.

[12] ——, "Dynamical consensus seeking of heterogeneous multiagent systems under input delays," *International Journal of Communication Systems*, vol. 26, no. 10, p. 12431258, 2013.

[13] Y. Zheng, J. Ma, and L. Wang, "Consensus of hybrid multi-agent systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1359–1365, 2018.

[14] D. Cartwright and F. Harary, "Structural balance: a generalization of heider's theory." *Psychological review*, vol. 63, no. 5, p. 277, 1956.

[15] R. Hegselmann, U. Krause *et al.*, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of artificial societies and social simulation*, vol. 5, no. 3, 2002.

[16] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.

[17] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2013.

[18] M. Yampolskiy, Y. Vorobeychik, X. D. Koutsoukos, P. Horvath, H. J. Leblanc, and J. Sztipanovits, "Resilient distributed consensus for tree topology," in *International Conference on High Confidence Networked Systems*, 2014, pp. 41–48.

[19] H. Zhang and S. Sundaram, "Robustness of information diffusion algorithms to locally bounded adversaries," in *American Control Conference*, 2011, pp. 5855–5861.

[20] H. J. Leblanc, H. Zhang, S. Sundaram, and X. Koutsoukos, "Resilient continuous-time consensus in fractional robust networks," in *American Control Conference*, 2013, pp. 1237–1242.

[21] S. M. Dibaji and H. Ishii, "Resilient consensus of double-integrator multi-agent systems," in *American Control Conference*, 2014, pp. 5139–5144.

[22] ——, "Consensus of second-order multi-agent systems in the presence of locally bounded faults," *Systems & Control Letters*, vol. 79, pp. 23–29, 2015.

[23] Y. Wu, X. He, and S. Liu, "Resilient consensus for multi-agent systems with quantized communication," in *American Control Conference*, 2016, pp. 5136–5140.

[24] Y. Wu and X. He, "Secure consensus control for multi-agent systems with attacks and communication delays," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 1, pp. 136–142, 2017.

[25] S. M. Dibaji, H. Ishii, and R. Tempo, "Resilient randomized quantized consensus with delayed information," in *IEEE 55th Conference on Decision and Control*, 2016, pp. 3505–3510.

[26] E. M. Shahrivar and S. Sundaram, "The game-theoretic formation of interconnections between networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 341–352, 2017.

[27] W. Abbas, A. Laszka, and X. Koutsoukos, "Improving network connectivity and robustness using trusted nodes with application to resilient consensus," *IEEE Transactions on Control of Network Systems*, p. DOI: 10.1109/TCNS.2017.2782486, 2017.

[28] D. Meng, Y. Jia, and J. Du, "Finite-time consensus for multiagent systems with cooperative and antagonistic interactions," *IEEE Trans Neural Netw Learn Syst*, vol. 27, no. 4, pp. 762–770, 2016.

[29] Z. Meng, G. Shi, K. H. Johansson, M. Cao, and Y. Hong, "Behaviors of networks with antagonistic interactions and switching topologies," *Automatica*, vol. 73, no. C, pp. 110–116, 2016.

[30] E. Seneta, *Non-negative matrices and Markov chains*. Springer Science & Business Media, 2006.

[31] J. Wolfowitz, "Products of indecomposable, aperiodic, stochastic matrices," *Proceedings of the American Mathematical Society*, vol. 14, no. 5, pp. 733–737, 1963.

[32] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1990.

[33] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Systems*, vol. 27, no. 2, pp. 71–82, 2007.