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

# Game of the Byzantine Generals on Time-Varying Graphs

**Yuke Li** * **Changbin Yu** **

* *School of Engineering, Westlake Institute for Advanced Study, Hangzhou, China 310024*
** *School of Engineering, Westlake Institute for Advanced Study, Hangzhou, China 310024*

**Abstract:** In this paper, we propose a game of the Byzantine Generals, which is a coordination game of agents seeking consensus by strategically transmitting information on a sequence of time-varying communication graphs. The first scenario of the game is where the generals cannot communicate with others at the same "level" in the communication graph. The second scenario is where those generals can. In either scenario, we examine the influences of the number of traitors and the decision rule held by the generals on equilibrium predictions of the game.

*Keywords:* Byzantine generals, distributed system, game theory, time-varying graphs, consensus-seeking, fault tolerance

## 1. INTRODUCTION

The Byzantine Generals Problem is a classic research question in the field of distributed systems (e.g., Lamport et al. (1982); Fischer (1983)). Fundamental to the functioning of those systems is a consensus reached by its components, possibly after information transmission among them. In the problem, several divisions of the Byzantine army are camped outside an enemy city, with each division commanded by a lieutenant. The lieutenants receive a message in the form of "attack" (0) or "not attack" (1) from a commander, send these messages among themselves, and decide on whether to attack via a majority voting. However, some of these generals, including the commander, might be traitorous, trying to prevent a good decision from being reached, such as by sending conflicting messages to others – a problem known as the "Byzantine fault." Lamport et al. (1982) has proposed an algorithm to ensure that the loyal generals will reach a consensus. The authors also show that the loyal generals can reach a consensus agreement if the fraction of traitorous generals is less than $\frac{1}{3}$. If the fraction is above $\frac{1}{3}$, consensus can no longer be guaranteed. Recently, Byzantine fault tolerant algorithms have been explored in a wide range of applications, such as in operation systems and signal processing (Duan et al, 2014; Kotla et al, 2007; Kailkhura et al, 2018; Sousa et al, 2018; Kapitza et al, 2012; Cowling et al, 2006).

In this paper, we propose to reformulate the original Byzantine Generals problem into a game-theoretic problem. In the game, the generals would transmit messages among themselves on a sequence of communication graphs. The strategic communication protocol is that the commander sends its messages to some lieutenants, who would decide on the messages sent to their "neighbors" in the communication graph (who are lieutenants at the same "level" and the next "level") and on their positions based on the messages received. At the end of the protocol, they will deliver a vote on their positions, which will determine their payoffs from the game. We will study the Nash equilibrium strategies of the generals – their strategic, mutual best responses. A significant difference from Lamport et al. (1982) is that we will still assume that the traitorous generals can send anything they wish. But the strategies of the traitors need to be optimally calibrated with those of the generals they communicate with and, importantly, vice versa. We will explore conditions for the game including the decision rule for the loyal generals and the sequence of communication graphs, under which different kinds of equilibria, an optimal consensus, a less optimal consensus or no consensus, may arise. As in Lamport et al. (1982), we will also address the question of "fault tolerance" – the maximum number of traitors in the game beyond which there will no longer be an optimal-consensus type of equilibria.

To the best of our knowledge, there are not many existing game-theoretic studies motivated by the Byzantine Generals problem. Halpern (2003) contains an informal discussion of how the problem can be thought of as a game between two opposing teams. The most relevant work, such as the electronic mail game formulated in Rubinstein (1989) as well as its follow-up work (e.g., Binmore and Samuelson (2001)), has shown that a lack of common knowledge generated by faulty communication can make coordinated actions impossible. However, an electronic mail game is usually played by two agents. Especially, the game does not cover the first type of Byzantine fault – the agents may send conflicting information to others, which is, instead, the focus of our study.

⋆ The first author thanks A. Stephen Morse (supported by the US Army Research Office Grant No.W911NF1910035) for helpful discussions initiating this work. The authors are supported by the National Natural Science Foundation of China (Grant No. 61761136005), and the Australian Research Council through DP-160104500 and DP- 190100887.

At a general level, our study relates to the literature on games of strategic information transmission. This strand of research started with the seminal work of Crawford and Sobel (1982), which develops a two-stage game with a sender transmitting a message to a receiver. The sender has private information about the statistical distribution of the state of the world unknown to the receiver and may send false messages to mislead the receiver if their interests do not align. The Crawford-Sobel model was later expanded to accommodate settings where a sender interacts with multiple receivers (Battaglini and Makarov, 2014; Golosov et al., 2014) or where multiple senders and receivers communicate on a static network (Galeotti et al., 2013). The contribution of our work is that we will consider the agents pass the messages on a dynamic sequence of communication graphs and that many of these agents are both a sender and a receiver.

Our study also relates to the literature on voting games, though to a lesser extent (e.g., Palfrey (1988); Downs et al. (1957); Davis et al. (1970); Baron and Ferejohn (1989); Austen-Smith and Banks (1988); Shepsle (1979); Myerson (2000); Snyder Jr et al. (2005); Riker (1982); Shepsle (2012); Coughlin (1992); Enelow and Hinich (1984)). As far as we know, a voting game normally does not cover a communication protocol between politicians and voters. Yet we do not consider the issues of legislative bargaining or party competitions or assume a certain distribution for the voters' preferences as in a voting game.

The paper will proceed as follows. First, we formulate the game of the Byzantine Generals. Second, we study two scenarios, with one where the lieutenants cannot communicate with others at the same "level" in the communication graph, and the other where they can. In the meantime, we explore equilibrium predictions in either scenario by assuming a specific decision rule. Third, we discuss future directions for this work.

## 2. PROBLEM FORMULATION

In the game of interest, there is a set of generals $\mathbf{n} = \{1, 2, \ldots, n\}$ connected on a communication graph $\mathbb{G} = \{\mathcal{V}, \mathcal{E}\}$, which is a directed graph. Any two connected agents $i$ and $j$ in $\mathcal{V}$ may have a directed edge $(i, j)$ or two directed edges $(i, j)$ and $(j, i)$ between them.

We partition $\mathbf{n}$ into $K \in \mathbb{N}$ disjoint subsets of generals,

$$\bigcup_{k=0}^{K} \mathbf{n}_k = \mathbf{n}. \tag{1}$$

The node set $\mathbf{n}_0 = \{u\}$ has $u$ designated as the commander. The node set $\mathbf{n}_k$ represents the set of the generals at level $k \in \{0, 1, \ldots, K\}$. For general $i$, let its neighbor set at level $k$ be $\mathcal{N}_{ik} = \{j : j \in \mathbf{n}_k \text{ and } (i, j) \in \mathcal{E} \text{ or } (j, i) \in \mathcal{E}\}$. We require that for $i \in \mathbf{n}_k$, $j \in \mathcal{N}_{ik+1}$ and $k \in \{0, 1, \ldots, K\}$, only $(i, j)$ exists in $\mathcal{E}$.

**Assumption 1: Message Passing.** At time $k \in \{0, 1, \ldots, K\}$, general $i$ at level $k$ send a message of 0 or 1 to each of their neighbors at levels $k$ and $k + 1$.

Assume that $i$ will send messages to his neighbors at level $k$ based on the messages received from his neighbors at level $k - 1$. $i$ will send messages to his neighbors at level $k + 1$ based on the messages received from his neighbors at both level $k - 1$ and level $k$.

In other words, for $i \in \mathbf{n}_k$, its messages to neighbors at level $k$ are denoted as the $|\mathcal{N}_{ik}|$-vector $[m_{ij}]_{1\times|\mathcal{N}_{ik}|}$ and determined by the map,

$$[m_{pi}] \mapsto [m_{ij}], \tag{2}$$

where

$$p \in \mathcal{N}_{ik-1} \text{ and } j \in \mathcal{N}_{ik}. \tag{3}$$

Its messages to neighbors at level $k+1$ are denoted as the $|\mathcal{N}_{ik+1}|$-vector $[m_{iq}]_{1\times|\mathcal{N}_{ik+1}|}$ and determined by the map,

$$[m_{pi}] \times [m_{ji}] \mapsto [m_{iq}], \tag{4}$$

where

$$p \in \mathcal{N}_{ik-1}, j \in \mathcal{N}_{ik} \text{ and } q \in \mathcal{N}_{ik+1}. \tag{5}$$

The messages are passed on a sequence of communication graphs $\mathbb{G}(k); k \in \{0, 1, \ldots, K\}$, each of which is a weakly connected digraph. We call this sequence an *ascending chain* of $\mathbb{G}$'s spanning subgraphs,

$$\mathbb{G}(k) \subset \mathbb{G}(k + 1), \tag{6}$$

which will reach $\mathbb{G}$ at time $K$,

$$\mathbb{G}(K) = \mathbb{G}, \tag{7}$$

with the following property

$$\mathcal{E}(k) \subset \mathcal{E}(k + 1) \tag{8}$$

where

$$\mathcal{E}(k+1) - \mathcal{E}(k) = \{(j, h) : j \in \mathbf{n}_{k+1} \text{ and } h \in \mathbf{n}_{k+1} \cup \mathbf{n}_{k+2}\} \tag{9}$$

and

$$\mathcal{E}(K) - \mathcal{E}(K - 1) = \{\{j, h\} : j, h \in \mathbf{n}_k\}. \tag{10}$$

The message $m_{jh} \in \{0, 1\}$ can thus be regarded as being passed on the directed edge $(j, h)$ in $\mathbb{G}(k + 1)$.

**Assumption 2: Own Position** There are two types of generals, loyal generals, and traitors, with their types known to only themselves. Each general $i$ needs to determine his own position $v_i \in \{0, 1\}$ based on messages received.

At level 0, the commander $u$ determines his position $v_u \in \{0, 1\}$. At level $k \in \{1, 2, \ldots, K\}$, the loyal generals determine their positions with the messages received from the generals at level $k - 1$ and level $k$, and send those positions to its neighbors at level $k + 1$. A loyal general's position is the value he sends to his neighbors at level $k+1$. By contrast, the value a traitor sends to his neighbors may not necessarily be his own position. A traitor may send whatever messages they would like to any neighbor.

**Assumption 2.1. Decision Rule** Each loyal general $i \in \mathbf{n}_k; k \in \{0, 1, \ldots, K\}$ can only observe messages sent to him from layers $k - 1$ and $k$. General $i$ will send the simple majority value of the messages received from layer $k - 1$ to his neighbors at layer $k$. He will adopt the simple majority value among the messages received from layer $k - 1$ and $k$ as his own position $v_i$, and send this value to the neighbors at layer $k + 1$. If "0"s and "1"s respectively make up one-half of all the received values, assume that $i$ may take either value as $v_i$. A traitor may send to his neighbors whatever he wishes.

**Assumption 3: Final Vote.** The game will end at time $K$. The *state* of the game is realized through a final,

majority vote on the positions of the generals described by vector

$$v = [v_i]_{1\times n}. \tag{11}$$

The loyal generals will vote truthfully on their positions, while the traitors may vote strategically. By "majority" is meant any desirable majority – in other words, at least a simple majority.

**Assumption 4: Possible Outcomes.** There are three possible types of outcomes. One outcome is no consensus, which is entirely possible when the total number of generals is even. The second and third outcomes are that a consensus on either 0 or 1 will be reached. Only one of the consensus outcomes can be regarded as the "general will," which we call the *optimal consensus*, and to which the traitors are opposed.

Assume for simplicity the preference structure of the loyal generals and traitors as follows. If no consensus is made, each general receives payoff 0. If they have an optimal consensus, each traitor receives payoff 1, and each loyal general receives payoff 2. If they have the less optimal consensus, each traitor receives payoff 2, and each loyal general receives payoff 1. The outcome and the payoffs are only realized at the end of the game. Other than the commander, the lieutenants have no prior information of whether 0 or 1 will be the optimal consensus.

In this game, it is natural to investigate the *subgame perfect Nash equilibrium*. Generals' strategies

$$m_{ij}, v_i \in \{0,1\} : i \in \mathbf{n}_k \text{ and } j \in \mathbf{n}_k \cup \mathbf{n}_{k+1}, 0 \le k \le K \tag{12}$$

are a subgame perfect Nash equilibrium of the game if for any general $i$ at any level $k$ unilaterally deviating to play an alternative strategy

$$m^*_{ij}, v^*_i \in \{0,1\}; j \in \mathbf{n}_k \cup \mathbf{n}_{k+1}, \tag{13}$$

he cannot receive a better payoff.

## 3. ANALYSIS AND RESULTS

### *3.1 Lieutenants Cannot Communicate*

We first consider the scenario where the lieutenants at each layer cannot communicate among themselves with Assumptions 1, 2.1, 3, and 4 in Section II.

The simplest game in this scenario only has two stages. Given that the lieutenants cannot communicate, the communication graph is a one-layer tree. In the set of generals $\mathbf{n} = \{1, 2, \ldots, n\}$, $u \in \mathbf{n}$ is the commander. At the root of the tree he sends a message, $m_{ui} \in \{0,1\}$, to each of the other $n-1$ generals, $i$, at layer 1, $i \in \mathbf{n} - \{u\}$. General $i$ will then choose his position $v_i \in \{0,1\}$ based on $m_{ui}$. Then a final majority vote will be delivered on $[v_i]_{1\times n}$ and the payoffs from the voting outcome are realized.

*Theorem 1.* If the following holds,

**1)** Only the commander is traitorous;
**2)** The game only has two stages and takes place on a tree, with the generals at layer 1 unable to communicate among themselves.

then any subgame perfect Nash equilibrium will not realize the general will, but the less optimal consensus.

*Proof of Theorem 1:* Assume without loss of generality that the general will is to realize "1" as the outcome, and that the traitor prefers "0" over "1". If a simple majority decides the final vote, there are two cases to consider.

**1)** The total number of generals is odd ($2m+1$, $m \in \mathbb{N}$). In this case, it is impossible not to reach a consensus, though a non-consensus might be the commander's most preferred outcome. Then one dominant strategy of $u$ is that $m_{ui} = 0$, $i \in \mathbf{n} - u$. In fact, in any equilibrium he only needs to send "0" to at least $m+1$ lieutenants.

By Decision Rule, a lieutenant at layer 1 will take whatever message from $u$ as his position. The above scenario is a subgame perfect Nash equilibrium because $u$ has no incentives to change its strategy, and any unilateral deviation by any single lieutenant from "0" to "1" is unable to change the outcome. If $u$ only sends "0" to $m$ lieutenants instead, there always exists one of those $m$ lieutenants preferring to deviate to adopt "1" as his position.

**2)** The total number of generals is even ($2m$, $m \in \mathbb{N}_+$). If the commander's most preferred outcome is "0", it will send "0" to at least $m$ lieutenants as above, and this will be a subgame perfect Nash equilibrium.

Even if the commander's most preferred outcome is a non-consensus, his best strategy is to opt for his second-best option, which is to realize "0" as the consensus outcome. For a non-consensus to occur, $u$ has to send exactly "0" to $m-1$ lieutenants and "1" to $m$ lieutenants, and use "0" as his position. This scenario cannot be a subgame perfect Nash equilibrium, because there will always be a lieutenant receiving "0" as the message defecting to adopt "1" as his position [1].

Then in any equilibrium, $u$ has to send "0" to at least $m$ lieutenants and takes "0" as his position, which means that he "sincerely" votes according to his preference and therefore will not defect. Also, none of the lieutenants will unilaterally deviate to take "1" as their positions, which either will not change the outcome, or will bring about a non-consensus.

In all, the traitorous commander can always exploit Decision Rule to disrupt the general will from being reached in equilibrium. A similar argument applies to the cases where the final vote is based on other kinds of majorities. □

A more general scenario than the one in Theorem 1 is that the game has multiple stages and the lieutenants at each layer cannot communicate among themselves. In the set of generals $\mathbf{n} = \{1, 2, \ldots, n\}$, $u \in \mathbf{n}$ is the commander sending a message, $m_{ui} \in \{0,1\}$, to each of the generals at layer 1, $i \in \mathbf{n}_1$. At layer $k \in \{1, 2, \ldots, K\}$, general $j$ will determine his position $v_i \in \{0,1\}$ based on the received messages from layer $k-1$ and send messages to his neighbors at layer $k+1$. Then a final majority vote will be delivered on $[v_i]_{1\times n}$ and the payoffs from the voting outcome are realized.

We first require that this game takes place on a multi-layer tree, which means that each non-commander general only has one sender from the previous layer and multiple receivers from the next layer.

[1] This reasoning applies in all other results in the paper and will not be repeated.

*Definition 1.* ("Critical Mass"). Suppose the total number of generals is $2m+1$ (or $2m$, $m \in \mathbb{N}$). For a subset of generals $\mathbf{n}_s$, if

**1)** They have the same most preferred outcome, "0" or "1" or nonconsensus.
**2)** As a result of their message passing, either the total number of generals who will take their favourable positions by Decision Rule is a desirable majority, or a nonconsensus will occur.

we say $\mathbf{n}_s$ is a "critical mass".

*Remark 1.* Given general $i$ at layer $k \in \{0, 1, 2, \ldots, K\}$ of the tree-structured communication graph $\mathbb{G}(k)$, denote the subtree of size $K-k+1$ with the root node $i$ as $T_i$. Let the set of nodes in the subtree $T_i$ be $\mathbf{n}_{T_i}$. When the vote is decided by a simple majority, a necessary and sufficient condition for a set of non-commander traitors $\mathbf{n}'$ to be a "critical mass" is the following inequality

$$(|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| > m{+}1 \text{ (or } m)) \vee (|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| = |\mathbf{n}'| = m{+}1 \text{ (or } m)) \tag{14}$$

*Theorem 2.* If the following holds,

**1)** Some generals may be traitorous.
**2)** The game has multiple stages, with the lieutenants at layer $k \in \{1, 2, \ldots, K\}$ receiving messages only from layer $k-1$ and transmitting messages to its neighbors on layer $k{+}1$, and the communication graph $\mathbb{G}(k)$ being a tree.
**3)** Assume that a simple majority will decide the final vote. The inequality in Remark 1 holds (does not hold) for the number of traitors.

Then there will not be (will be) a subgame perfect Nash equilibrium realizing the general will.

*Proof of Theorem 2:* As in Remark 1, let the set of traitors be $\mathbf{n}' \in \mathbf{n}$. We first consider the case when $|\mathbf{n}| = 2m+1$, and proceed to prove the condition in Remark 1 to be both necessary and sufficient.

Necessity. There are two cases to consider.

**1)** If $|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| = m{+}1$ but $\sum_{i\in\mathbf{n}'} |\mathbf{n}_{T_i}| \neq |\mathbf{n}'|$, it means that some of these $m+1$ generals are loyal because $|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| \geq |\mathbf{n}'|$. It cannot be an equilibrium that these $m+1$ generals vote "0" and the other $m$ generals vote "1," from which a unilateral deviation by any loyal general among the $m+1$ generals will be profitable.
**2)** If $|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| < m{+}1$, the number of generals holding 0 as their positions will not be a majority because the commander is loyal. The only kind of subgame perfect Nash equilibrium in the game is 1 realized as the outcome.

Sufficiency. When $|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| > m{+}1$, any traitor $i$ in any equilibrium will only send 0 to its neighbors in the next layer. By Decision Rule, the number of generals holding 0 as their positions will be a simple majority. Then the only kind of subgame perfect Nash equilibrium in the game is 0 realized as the outcome.

Alternatively, when $|\bigcup_{i\in\mathbf{n}'} \mathbf{n}_{T_i}| = m+1$, which is the second inequality in the condition in Remark 1. It is a subgame perfect Nash equilibrium that these $m{+}1$ traitors vote "0" and the $m$ loyal generals vote "1", because none can unilaterally deviate to make himself better off.

When $|\mathbf{n}| = 2m$, or the final vote is decided by other kinds of majorities, a similar argument applies. □

Theorem 1 is a special case of Theorem 2 – if the commander is traitorous, the condition in Remark 1 easily holds. Theorems 1 and 2 could still apply when the loyal generals adopt a different decision rule – a loyal general may take the exact opposite value of the message received from the commander as his own position. Despite the existence of a traitorous commander, the general will might be enforced in equilibrium if the loyal generals have no specific decision rule. (e.g., acting randomly.)

We now consider the scenario in which the lieutenants at the same layer cannot communicate, but the communication graph may not be a tree. Given the neighborhood structure of the generals, it would be particularly difficult to generalize on a necessary and sufficient condition by which a critical mass will form. Therefore, we derive a sufficient condition for a subset of generals being a critical mass below.

*Theorem 3.* If the following holds,

**1)** Both the commander and the generals from level 1 are traitorous with the same preferred positions.
**2)** The game has multiple stages, with the lieutenants at layer $k \in \{1, 2, \ldots, K\}$ receiving messages only from layer $k-1$ and transmitting messages to its neighbors on layer $k+1$.

Then there will not be a subgame perfect Nash equilibrium realizing the general will.

*Proof of Theorem 3:* As before, suppose without loss of generality that the general will is "1". If the commander and the generals from layer 1 are traitorous, their goal is to have enough generals adopt their positions eventually.

At layer $k$, any traitorous general $i$ has a dominant strategy, which is to send at least $m_i \in \mathbb{N}$ neighbors at layer $k+1$ his real position, "0".

By Decision Rule, at layer $k+1$, a loyal general at layer $k{+}1$ will adopt the majority value of the received messages as their positions, and send his positions onto layer $k+2$. A traitorous general at layer $k+1$ will do the same as those traitors did at layer $k$.

None will unilaterally deviate because

**1)** the traitors have no incentives to change their strategies.
**2)** given the traitors' strategies, none of the loyal generals can unilaterally reverse the game outcome.

Then in any equilibrium, the traitors will make sure there is a majority to adopt their preferred positions. □

There could be more sufficient conditions along similar lines. For instance, the commander does not have to be traitorous. Alternatively, we could further require more generals below layer 1 to be traitorous. This leads to a more general discussion in Section 3.2 on the connection between the existence of a critical mass of traitors and the equilibrium prediction.

### *3.2 Lieutenants Can Communicate*

Now we consider the scenario where the lieutenants at each layer may communicate among themselves, with the same set of assumptions as in Section 3.1.

The primary difference from the second game is that at layer $k \in \{1, 2, \ldots, K\}$, general $i$ will determine his position $v_i \in \{0, 1\}$ and send messages to his neighbors at layer $k+1$ based on the messages from layer $k-1$ and $k$. Then when receiving conflicting messages, it may be impossible to deduce using Decision Rule whether the traitors are from his layer or the previous layer or both.

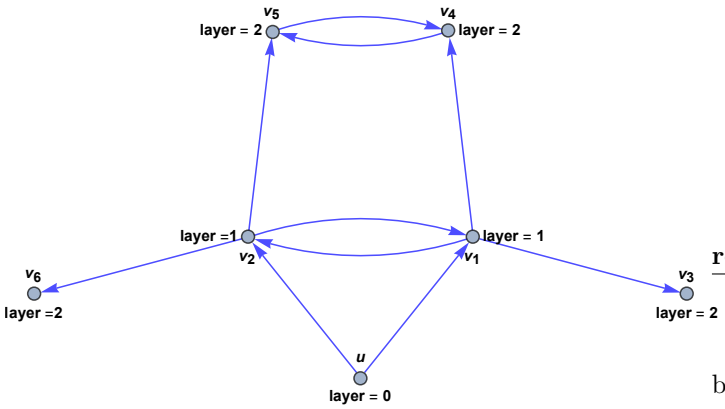


Fig. 1. Traitors exist from both layer 1 and layer 2

*Example 1.* Figure 1 shows a communication graph $\mathbb{G}(2)$ (with a commander and two layers of lieutenants). Suppose the second lieutenant $v_1$ from layer 1 and the fifth lieutenant $v_5$ from layer 2 are traitorous.

On receiving the likely identical messages from both $v_1$ and $v_5$, $v_4$ may not know both of them to be traitorous. On receiving the likely conflicting messages from the commander $u$ and $v_2$, $v_1$ may not tell whether one or both of them are traitorous.

The following statement is a tautology. If we have

**1)** Some generals are traitorous.
**2)** The game is played as described by Decision Rule.
**3)** The traitors constitute a critical mass

then there will not be a subgame perfect Nash equilibrium realizing the general will. Given a set of generals $\mathbf{n}$ and a sequence of communication graphs $\mathbb{G}(k), k \in \{0, 1, \ldots, K\}$, the existence conditions for a critical mass of traitors would be unrealistic to determine ex ante. It would be realistic, however, to examine *whether* such a critical mass exists based on an algorithm. And we propose one such algorithm below.

*Theorem 4.* By Algorithm 1, the set of traitors in the game $\mathbf{n}'$ are a critical mass if $\mathbf{n}_s$ is a desirable majority.

*Proof of Theorem 4:* By Algorithm 1, if general $i$ at layer $k \in \{0, 1, \ldots, K\}$ is a traitor, it is then an element of the set $\mathbf{n}_s$.

A loyal general $i$ at layer $k \in \{0, 1, \ldots, K\}$ first determines what messages to send to its neighbors at the same layer. If he has more traitors in his neighbors than loyal generals from layer $k-1$,

**Algorithm 1**

**Input:**
$\mathbf{n}_k, \mathbb{G}(k); k \in \{0, 1, \ldots, K\}, \mathbf{n}'$
**Output:**
$\mathbf{n}_s$
**Initialize:**
$\mathbf{n}_s \leftarrow \emptyset$, $\mathbf{n}_s^* \leftarrow \emptyset$, $\bar{\mathbf{n}}_s \leftarrow \emptyset$
**for** $0 \le k \le K$ **do**
  **for** $i \in \mathbf{n}_k$ **do**
    **if** $i \in \mathbf{n}'$ **then**
    $\mathbf{n}_s \leftarrow \mathbf{n}_s \cup \{i\}$
    **else if** $|\{j \in \mathbf{n}_{k-1} : \{i,j\} \in \mathcal{E}_k \text{ and } j \in \mathbf{n}'\}| >$
      $|\{j \in \mathbf{n}_{k-1} : \{i,j\} \in \mathcal{E}_k \text{ and } j \notin \mathbf{n}'\}|$ **then**
    $\mathbf{n}_s^* \leftarrow \mathbf{n}_s^* \cup \{i\}$
    **else** go to the next general in $\mathbf{n}_k$
  **for** $i \in \mathbf{n}_k$ **do**
    **if** $|\{j \in \mathbf{n}_{k-1} \cup \mathbf{n}_k : j \in \mathbf{n}_s \cup \mathbf{n}_s^* \text{ and } \{i,j\} \in \mathcal{E}_k\}| >$
      $|\{j \in \mathbf{n}_{k-1} \cup \mathbf{n}_k : j \notin \mathbf{n}_s \cup \mathbf{n}_s^* \text{ and } \{i,j\} \in \mathcal{E}_k\}|$ **then**
      $\bar{\mathbf{n}}_s \leftarrow \bar{\mathbf{n}}_s \cup \{i\}$
    **else** go to the next general in $\mathbf{n}_k$.
  $\mathbf{n}_s \leftarrow \mathbf{n}_s \cup \bar{\mathbf{n}}_s$
**return** $\mathbf{n}_s$

$$|\{j \in \mathbf{n}_{k-1} : \{i,j\} \in \mathcal{E}_k \text{ and } j \in \mathbf{n}'\}| > |\{j \in \mathbf{n}_{k-1} : \{i,j\} \in \mathcal{E}_k \text{ and } j \notin \mathbf{n}'\}| \quad (15)$$

by Decision Rule that he will send the traitors' preferred value to its neighbors at layer $k$ in the game. Then we let $i$ be an element of the set $\mathbf{n}_s^*$.

By Decision Rule, the loyal general $i$ will determine his position as well as the messages sent to his neighbors at layer $k+1$ based on messages received from layer $k-1$ and layer $k$.

Then if $i$ has more neighbors from layer $k-1$ and layer $k$ who will send him a traitor's preferred value than neighbors from these two layers who will not,

$$|\{j \in \mathbf{n}_{k-1} \cup \mathbf{n}_k : j \in \mathbf{n}_s \cup \mathbf{n}_s^* \text{ and } \{i,j\} \in \mathcal{E}_k\}| > |\{j \in \mathbf{n}_{k-1} \cup \mathbf{n}_k : j \notin \mathbf{n}_s \cup \mathbf{n}_s^* \text{ and } \{i,j\} \in \mathcal{E}_k\}| \quad (16)$$

he will adopt this value as his position and send it to neighbors at layer $k+1$. We let $i$ be an element of the set $\bar{\mathbf{n}}_s$.

At the end of the $k$-th iteration ($k \in \{0, 1, \ldots, K\}$), the updated set $\mathbf{n}_s$ denotes the union of the set of traitors and the set of loyal generals who will adopt the same positions with the traitors up to layer $k$.

At the end of the algorithm, if $|\mathbf{n}_s|$ is a desirable majority, then the original set of traitors $\mathbf{n}'$ are a critical mass by definition. □

Finding the minimum number of traitors who will constitute a critical mass is equivalent to solving the following integer programming problem for the communication graph $\mathbb{G}$.

$$\min |\mathbf{n}'| \quad \text{such that } \mathbf{n}' \subset \mathbf{n}, \ |\mathbf{n}_s| \text{ is a desirable majority} \quad (17)$$

The integer programming problem is unlikely to be computed in polynomial time. Nevertheless, it does not have to be computed – If the commander is traitorous, his strategy

[2] $\mathbf{n}_s$ is obtained by Algorithm 1.

is to selectively influence those from below layers such that a desirable majority holding his preferred positions will eventually be obtained, regardless of whether other traitors might exist in the game. He could revise his strategy accordingly if such a goal falls short. Then in any subgame perfect Nash equilibrium, the general will is not realized with the existence of a traitorous commander.

By this reasoning, if the commander has a preferred value (instead of a nonconsensus), he will have a majority voting for this value in equilibrium. Therefore, the minimum number of traitors that may constitute a critical mass for all instances of the above problem is 1.

## 4. CONCLUSION

In this paper, we have formulated a game of the Byzantine generals on a sequence of time-varying communication graphs. We have proposed a decision rule and studied equilibrium predictions of the game with the presence of traitorous generals. We found that when a traitorous commander exists, there will be no subgame perfect Nash equilibrium realizing the general will under either rule. In other cases, the number of traitors will impact equilibrium predictions.

For future work, we would like to investigate the more general case of the game, where a lieutenant can act at multiple times in the game. It will be worthy of studying how a general may deduce who are the traitors from the messages transmitted, as well as how the decision rules may influence the game equilibria.

## REFERENCES

Austen-Smith, D. and Banks, J. (1988). Elections, coalitions, and legislative outcomes. *American Political Science Review*, 82(2), 405–422.

Baron, D.P. and Ferejohn, J.A. (1989). Bargaining in legislatures. *American Political Science Review*, 83(4), 1181–1206.

Battaglini, M. and Makarov, U. (2014). Cheap talk with multiple audiences: An experimental analysis. *Games and Economic Behavior*, 83, 147–164.

Binmore, K. and Samuelson, L. (2001). Coordinated action in the electronic mail game. *Games and Economic Behavior*, 35(1-2), 6–30.

Coughlin, P.J. (1992). *Probabilistic Voting Theory*. Cambridge University Press.

Cowling, James and Myers, Daniel and Liskov, Barbara and Rodrigues, Rodrigo and Shrira, Liuba (2006). HQ replication: A hybrid quorum protocol for Byzantine fault tolerance. *Proceedings of the 7th Symposium on Operating Systems Design and Implementation,* 177–190, 2006.

Crawford, V.P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, 1431–1451.

Davis, O.A., Hinich, M.J., and Ordeshook, P.C. (1970). An expository development of a mathematical model of the electoral process. *American Political Science Review*, 64(2), 426–448.

Downs, A. et al. (1957). An Economic Theory of Democracy. New York: Harper, 1957.

Duan, Sisi and Peisert, Sean and Levitt, Karl N. (2014). hBFT: speculative Byzantine fault tolerance with minimum cost. *IEEE Transactions on Dependable and Secure Computing*, 12(1), 58–70, 2014.

Enelow, J.M. and Hinich, M.J. (1984). *The Spatial Theory of Voting: An Introduction*. CUP Archive.

Fischer, M.J. (1983). The consensus problem in unreliable distributed systems (a brief survey). In *International Conference on Fundamentals of Computation Theory*, 127–140. Springer.

Galeotti, A., Ghiglino, C., and Squintani, F. (2013). Strategic information transmission networks. *Journal of Economic Theory*, 148(5), 1751–1769.

Golosov, M., Skreta, V., Tsyvinski, A., and Wilson, A. (2014). Dynamic strategic information transmission. *Journal of Economic Theory*, 151, 304–341.

Halpern, J.Y. (2003). A computer scientist's look at game theory. *Games and Economic Behavior*, 45(1), 114–131.

Kapitza, Rdiger and Behl, Johannes and Cachin, Christian and Distler, Tobias and Kuhnle, Simon and Mohammadi, Seyed Vahid and Schrder-Preikschat, Wolfgang and Stengel, Klaus (2012). CheapBFT: resource-efficient byzantine fault tolerance. *Proceedings of the 7th ACM European Conference on Computer Systems*, 295–308, 2012.

Kailkhura, Bhavya and Vempaty, Aditya and Varshney, Pramod K (2018). Collaborative Spectrum Sensing in the Presence of Byzantine Attacks. *Cooperative and Graph Signal Processing*, 505–522, 2018.

Kotla, Ramakrishna and Alvisi, Lorenzo and Dahlin, Mike and Clement, Allen and Wong, Edmund (2007). Zyzzyva: speculative byzantine fault tolerance. *ACM SIGOPS Operating Systems Review*, 41(6), 45–58, 2007.

Lamport, L., Shostak, R., and Pease, M. (1982). The Byzantine generals problem. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 4(3), 382–401.

Myerson, R.B. (2000). Large Poisson games. *Journal of Economic Theory*, 94(1), 7–45.

Palfrey, T.R. (1988). A Mathematical Proof of Duverger's Law. Social Science Working Paper, 688. California Institute of Technology. Pasadena, CA.

Riker, W.H. (1982). The two-party system and Duverger's law: an essay on the history of political science. *American Political Science Review*, 76(4), 753–766.

Rubinstein, A. (1989). The electronic mail game: strategic behavior under "almost common knowledge". *The American Economic Review*, 385–391.

Shepsle, K. (2012). *Models of Multiparty Electoral Competition*. Routledge.

Shepsle, K.A. (1979). Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science*, 27–59.

Snyder Jr, J.M., Ting, M.M., and Ansolabehere, S. (2005). Legislative bargaining under weighted voting. *American Economic Review*, 95(4), 981–1004.

Sousa, Joao and Bessani, Alysson and Vukolic, Marko (2018). A byzantine fault-tolerant ordering service for the hyperledger fabric blockchain platform. *The 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 51–58, 2018.