

Gracefully Degrading Fair Exchange with Security Modules

Gildas Avoine¹, Felix Gärtner², Rachid Guerraoui³, and Marko Vukolić³

¹ Security and Cryptography Laboratory, EPFL, Switzerland

² Dependable Distributed Systems Laboratory, RWTH Aachen University, Germany

³ Distributed Programming Laboratory, EPFL, Switzerland

Abstract. The *fair exchange* problem is key to trading electronic items in systems of mutually untrusted parties. In modern variants of such systems, each party is equipped with a security module. The security modules trust each other but can only communicate by exchanging messages through their untrusted host parties, that could drop those messages.

We describe a synchronous algorithm that ensures deterministic *fair exchange* if a majority of parties are honest, which is optimal in terms of resilience. If there is no honest majority, our algorithm degrades gracefully: it ensures that the probability of unfairness can be made arbitrarily low.

Our algorithm uses, as an underlying building block, an early-stopping subprotocol that solves, in a general omission failure model, a specific variant of consensus we call *biased consensus*. Interestingly, this modular approach combines concepts from both cryptography and distributed computing, to derive new results on the classical fair exchange problem.

1 Introduction

1.1 Motivation

Fair exchange (see e.g. [5–7, 10–12, 14, 31]) is a fundamental problem in systems with electronic business transactions. In fair exchange, the participating parties start with an item they want to trade for another item. They possess an executable description of the desired item, typically a boolean function with which an arbitrary item can be checked for the desired properties. Furthermore, they know from which party to expect the desired item and which party is expecting their own item. An algorithm that solves fair exchange must ensure that every honest party eventually either delivers its desired item or aborts the exchange (*termination* property). The abort option however is excluded if no party misbehaves and all items match their descriptions (*effectiveness* property). The algorithm should also guarantee that, if the desired item of any party does not match its description, then no party can obtain any (useful) information about any other item (*fairness* property).

Fair exchange is easily solvable using a *trusted third party* through which all items can be exchanged [13]. The involvement of the trusted third party can be

reduced using *optimistic* schemes where participation of the trusted third party is only necessary if something goes wrong [1]. The context of this paper is one where the trusted third party is a *virtual* entity, distributed within all untrusted parties, as we explain below.

We consider in this paper a system where each party hosts a *security module* that is *tamper proof* (Fig. 1). Recently, manufacturers have begun to equip hardware with such modules: these include for instance the “Embedded Security Subsystem” within the recent IBM Thinkpad, or the IBM 4758 secure co-processor board [15]. In fact, a large body of computer and device manufacturers has founded the Trusted Computing Group (TCG) [32] to promote this idea. Besides security modules being tamper proof, the software running within the security modules is certified and they can communicate through secure channels. In certain settings, the overall system can even assumed to be synchronous, i.e., it is reasonable to assume an upper bound on the relative speeds of honest parties (and their security modules) as well as on the communication delays between them. However, dishonest parties can still drop messages exchanged between the underlying security modules in order to violate the fairness of the exchange in their favor, i.e., obtain an item without giving away their own.

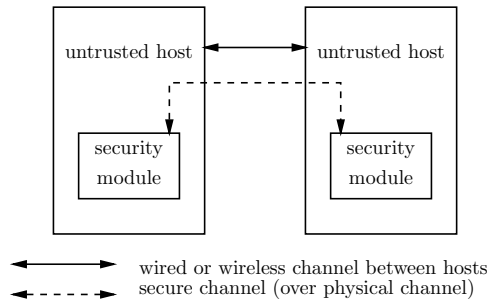


Fig. 1. Hosts and security modules.

The contribution of this paper is a synchronous distributed algorithm aimed at exchanging electronic items among multiple untrusted parties, each hosting a security module. The algorithm provides the two following complementary features:

1. If a majority of parties is honest, then the algorithm deterministically guarantees the termination, effectiveness and fairness properties of fair exchange. This is optimal in terms of resilience: we indeed show that, even in a synchronous model with security modules, no deterministic algorithm solves fair exchange if half of the parties are dishonest.
2. If at least half of the parties turn out to be dishonest, then our algorithm degrades gracefully in the following sense. It still guarantees the termination and effectiveness properties of fair exchange, as well as ensures that the

probability of violating fairness can be made arbitrarily low. We supply the probability distribution that optimizes the average complexity of the algorithm, in terms of its number of communication rounds, and we show that the probability of violating fairness is inversely proportional to the average algorithm complexity.

Our algorithm is made of three phases, and we give the intuition underlying each phase below.

1. In the first phase, which we call the *initialization* phase, the security modules exchange the items that are supposed to be traded by their untrusted hosts. These items are not delivered by the security modules to their untrusted hosts: this is only performed if the third phase (below) terminates *successfully*. Any security module can decide here to abort the exchange if some item is missing or does not match its expected description. The security module hosted by the party that initiates the exchange also selects here a *random number* k that it disseminates to all other security modules. The role of this random number is crucial in the second phase of the algorithm.
2. In the second phase, which we call the *fake* phase, all security modules exchange messages during k rounds; each round following the same communication pattern as in the third phase (below). The fact that the random number k , determined in the first phase, is not accessible to the untrusted parties is fundamental here. Roughly speaking, the goal of the *fake* phase is to make the probability, for *any* number of dishonest parties to successfully guess when the actual agreement phase is taking place (third phase below), arbitrarily low. If any dishonest party drops a message towards a honest party in this *fake* phase, the security module hosted by the latter simply aborts the exchange and forces other modules to abort the exchange as well, thus penalizing any dishonest host that might try to bias the exchange in its favor.
3. In the third phase, which we call the *agreement* phase, the security modules solve a problem we call *biased consensus*. In this problem, the processes (in our case the security modules) start from an initial binary value (a proposal) and need to decide on a final binary value: either to abort the exchange or commit it (and deliver the items to their untrusted hosts). Unlike in consensus [17], but like in NBAC (non-blocking atomic commit) [9,30], the problem is biased towards 0: no process can decide 1 if some process proposes 0 (to avoid trivial solutions, the processes are supposed to decide 1 if no process fails or proposes 0). The agreement aspect of this problem is however different from consensus and NBAC; we simply require here that, if some process decides 1, then no correct process decides 0. We consider an early stopping algorithm that solves this problem in a model with general omissions, along the lines of [28].

Underlying our main contribution, i.e., a new gracefully degrading fair exchange algorithm, we contribute in bridging the gap between security problems

(fair exchange) and traditional distributed computing problems (consensus-like problems). We show indeed that deterministic fair exchange in a model with security modules is equivalent to biased consensus. By proving that biased consensus is impossible in a synchronous model [24] with general omission failures [29] if half of the processes can be faulty, we establish a lower bound for fair exchange in a model with tamper proof modules.

1.2 Roadmap

Section 2 defines our system model. Section 3 recalls the fair exchange problem, introduces biased consensus, and shows their equivalence in a model with security modules. We also state the impossibility of deterministic fair exchange without a honest majority, which motivates our notion of gracefully degrading fair exchange. Section 4 describes our gracefully degrading fair exchange algorithm and states its correctness. Section 5 concludes the paper by discussing related work.

2 Model

The system we consider is composed of a set of processes, some modeling untrusted hosts and the other modeling security modules. These processes communicate by exchanging messages.

2.1 Untrusted hosts and security modules

More precisely, the set of processes we consider is divided into two disjoint classes: *untrusted hosts* (or simply *hosts*) and *security modules*. Two processes connected by a physical channel are said to be adjacent. We assume that there exists a fully connected communication topology between the hosts, i.e., any two hosts are adjacent. Furthermore, we assume that every host process P_A is adjacent to exactly one security module process G_A (i.e., there is a bijective mapping between security modules and hosts): we say that P_A is *associated with* G_A . No two security modules are adjacent. In other words, for any two security modules G_A and G_B to communicate, they need to do so through their hosts P_A and P_B . This indirection provides the abstraction of an overlay network at the level of security modules. We call the part of the system consisting of security modules, and the virtual communication links between them, the *security subsystem*. We also assume that every host (resp. security module) has the knowledge about the entire set of other hosts (resp. security modules) that participate in the protocol. We call the part of the system consisting of hosts and the communication links between them the *untrusted system*. The notion of association can be extended to systems, meaning that, for any given untrusted system, the *associated security subsystem* is the system consisting of all security modules associated to any host in that untrusted system.

2.2 Security modules and virtual channels

Security modules are interconnected by a virtual communication network with bidirectional channels over the physical communication network among the hosts. For simplicity, we denote the participants processes (the security modules) by G_1, \dots, G_n . We assume that between any two security modules G_i and G_j , the following properties are guaranteed: (1) Message contents remain secret from unauthorized entities; (2) If a message is delivered at G_j , then it was previously sent by G_i ; (3) Replayed messages are detected; (4) Message contents are not tampered with during transmission, i.e., any change during transmission will be detected and the message will be discarded; (5) If a message is sent by G_i to G_j and G_j is ready to receive the message, then the message will be delivered at G_j within some known bound Δ on the waiting time.

2.3 Trust and adversary model

Security modules can be trusted by other security modules or hosts, and hosts cannot be trusted by anybody. Hosts may be malicious, i.e., they may actively try to fool a protocol by not sending any message, sending wrong messages, or even sending the right messages at the wrong time. We assume however that hosts are computationally bounded, i.e., brute force attacks on secure channels are not possible. A malicious host may inhibit *all* communication between its associated security module and the outside world, yielding a channel in which messages can be lost.

A host *misbehaves* if it does not correctly follow the prescribed algorithm and we say that the host is *dishonest*. Otherwise it is said to be *honest*. Misbehavior is unrestricted (but computationally bounded as we pointed out). Security modules always follow their protocol, but since their associated hosts can inhibit all communication, this results in a system model of security modules with unreliable channels (the model of *general omission* [29], i.e., where messages may not be sent or received). In such systems, misbehavior (i.e., failing to send or receive a message) is sometimes termed *failure*. We call security modules associated with honest hosts *correct*, whereas those associated with dishonest hosts *faulty*. In a set of n hosts, we use t to denote a bound on the number of hosts which are allowed to misbehave and f the number of hosts which actually do misbehave ($f \leq t$). Sometimes we restrict our attention to the case where $t < n/2$, i.e., where a majority of hosts is assumed to be honest. We call this the *honest/correct majority assumption*.

Our model of the adversary is based on the strongest possible attack, the case in which all of the f dishonest hosts collude. We assume that adversary knows all the algorithms and probability distributions used.

3 Variations on Fair Exchange and Impossibility Results

In this section we recall the definition of fair exchange (FE), and we show that this problem, at the level of untrusted hosts, is in a precise sense equivalent to

a problem that we call *biased consensus* (BC), at the level of the underlying security modules. Then, we state that biased consensus is impossible if half of the processes can be faulty and derive the impossibility of fair exchange if half (or more) of the hosts are dishonest. This motivates our definition of a weaker variant of fair exchange, the gracefully degrading FE.

3.1 Fair exchange

Definition 1 (Fair Exchange). *An algorithm solves fair exchange (FE) if it satisfies the following properties [1, 27]*

- (Timeliness) *Every honest host eventually terminates.*
- (Effectiveness) *If no host misbehaves and if all items match their descriptions then, upon termination, every host has the expected item.*
- (Fairness) *If the desired item of any host does not match its description, or any honest host does not obtain any (useful) information about the expected item, then no host can obtain any (useful) information about any other host’s item.*

In case a host terminates without receiving the expected item, that host receives an abort indication (denoted \perp). The Timeliness property ensures that every honest host can be sure that at some point in time the algorithm will terminate. The Effectiveness property states what should happen if all goes well. Finally, the Fairness property postulates restrictions on the information flow for the case where something goes wrong in the protocol.⁴ Note that the first precondition of the Fairness property (“if the desired item of any host does not match the description...”) is very important. Without this condition, a “successful” outcome of the exchange would be possible even if an item does not match the expected description, which should clearly be considered unfair.

3.2 Biased Consensus

Consider the following variant of consensus, we call *biased consensus* in a model where processes can fail by general omissions [29].

Definition 2 (Biased Consensus). *An algorithm solves biased consensus (BC) if it satisfies the following properties:*

- (Termination) *Every correct process eventually decides.*
- (Non-Triviality) *If no process is faulty or proposes 0, then no correct process decides 0.*
- (Validity) *No process decides 1 if some process proposes 0.*
- (Biased Agreement) *If any process decides 1, then no correct process decides 0.*

⁴ We use here the concept of information flow to define fairness in a way that cleanly separates the distinct classes of *safety*, *liveness*, and *security* properties in the specification of the problem [26].

Processes invoke *biased consensus* using primitive $BCpropose(vote)$, $vote$ being a binary value, 0 or 1. Possible decisions are also 0 (*abort*) and 1 (*commit*). *Termination*, *Non-Triviality* and *Validity* are the same as in NBAC, whereas the *Biased Agreement* is weaker than the *Agreement* property of NBAC [30].

We show below that FE and BC are equivalent in our model.

```

FairExchange( $myitem, description, source, destination$ ) returns  $item$  {
   $\langle$ send  $myitem$  to  $destination$  over secure channel $\rangle$ 
  timed wait for  $\langle$ expected item  $i$  from  $source$  over secure channel $\rangle$ 
   $\langle$ check  $description$  on  $i$  $\rangle$ 
  if  $\langle$ check succeeds and no timeout $\rangle$ 
  then  $vote := 1$  else  $vote := 0$  endif
   $result := BCpropose(vote)$ 
  if  $result = 1$  then return  $i$  else return  $\langle$ abort $\rangle$  endif
}

```

Fig. 2. Using biased consensus to implement fair exchange: code of every host.

Theorem 1. *Biased consensus is solvable in the security subsystem, iff fair exchange is solvable in the associated untrusted system.*

Proof. (1) Assume that we have a solution to BC in the security subsystem consisting of security modules G_1, \dots, G_n . Now consider the algorithm depicted in Fig. 2. This is a wrapper around the BC solution that solves FE at the level of the hosts. In other words, it is a reduction of FE into BC in our model. In the algorithm, a host hands to the associated security module its item and the executable description of the desired item, as well as identifiers of the hosts with which items should be exchanged. The security module exchanges the item with its partners, then checks the received item (*initialization* phase). Finally all security modules agree on the outcome using BC (*agreement* phase). The proposal value for BC is 1 if the check was successful and no abort was requested by the host in the meantime. If BC terminates with the process deciding 1, then the security module releases the item to the host. We now discuss each of the properties of fair exchange. The *Timeliness* property of FE is guaranteed by the *Termination* property of BC and the synchronous model assumption. Consider *Effectiveness* and assume that all participating hosts are honest and all items match their descriptions. All votes for BC will be 1. Now the *Non-Triviality* property of BC guarantees that all processes (recall that every process is correct) will decide 1 and subsequently return the item to their hosts. Consider now *Fairness* and observe that, in our adversary model, no dishonest host can derive any useful information from merely observing messages exchanged over the secure channels. The only way to receive information is through the interface of the *FairExchange* procedure. If one item does not match the description at some process, then this process will engage in BC with a $vote = 0$ (note that

this will happen even if the associated host is dishonest). *Validity* of BC implies that the exchange results in no process deciding 1, so none of the hosts receives anything from the exchange. Additionally, if some honest host receives nothing through the exchange, then the *Biased Agreement* property of BC implies that no host can receive anything.

(2) Conversely, BC can be implemented using FE by invoking at every process G_i .⁵:

$$BCpropose(vote_i) = FairExchange(vote_i, 1, G_{i-1}, G_{i+1}).$$

Here we assume that if FE returns *abort* instead of the item, BC returns 0. So the votes, being exchange items in this case, are exchanged in a circular fashion among security modules. It is not difficult to see that FE properties guarantee the properties of BC. This is immediate for *Termination* and *Non-Triviality*. Consider now *Validity* and assume that G_j proposes 0 to BC. The item description checking at G_{j+1} will fail and the first part of FE *Fairness* (“If the desired item of any host does not match its description. . .”) guarantees that every process that decides in BC decides 0. The second part of *Fairness* guarantees *Biased Agreement*.⁶ \square

Theorem 2. *Consider a synchronous system where processes can fail by general omissions. No algorithm solves biased consensus if $\lceil \frac{n}{2} \rceil$ processes can be faulty.*

We omit the proof of the Theorem 2 due to the lack of space. The proof can be found in the full version of the paper [2].

A direct corollary of the Theorems 1 and 2 leads to derive the following result:

Theorem 3. *Consider our model of untrusted hosts and security modules. No algorithm solves fair exchange if half of the hosts can be dishonest.*

3.3 Early Stopping Biased Consensus Algorithm

The BC algorithm we give here (in Fig. 3) is an adaptation of the early-stopping synchronous consensus algorithm of [28]. This algorithm solves BC if there is a majority of correct processes ($t < n/2$). It is early stopping in the sense that every process terminates in at most $\min(f + 2, t + 1)$ rounds. There are mainly two differences with the consensus algorithm of [28]: (1) the processes agree on a *vector* of initially proposed values rather than on a *set* of those (in the case of consensus); (2) we also introduce *dummy* messages to have a *full information protocol* [24], i.e., to have a uniform communication pattern in every round. In other words, where in the original algorithm of [28] process G_i did not send a message to process G_j in round r , in our algorithm process G_i sends a *dummy*

⁵ To be precise, G_1 invokes $FairExchange(vote_1, 1, G_n, G_2)$ and G_n invokes $FairExchange(vote_n, 1, G_{n-1}, G_1)$.

⁶ Note that, in contrast to BC, FE satisfies an information-flow (i.e., security) property [26]. This is why it was necessary to argue about the special properties of security modules when reducing FE to BC and not vice versa.

BCpropose($vote_i$) returns *decision* {

```
1:  $Vote_i := \perp^n$ ;  $Vote_i[i] := vote_i$ ;  $new_i := \perp^n$ ;  $new_i[i] := vote_i$ ;
2:  $locked_i := \emptyset$ ;  $suspected_i := \emptyset$ ; % r=0 %
3: for  $r := 1, 2, \dots, t + 1$  do % r: round number %
4: begin_ round
5:   foreach  $p_j$  do
6:     if  $p_j \in suspected_i$  then  $dummy := 1$  else  $dummy := 0$  endif
7:     send ( $new_i, locked_i, dummy$ ) to  $G_j$ 
8:   enddo
9:    $new_i := \perp^n$ 
10:  foreach  $G_j \notin suspected_i$  do
11:    if ( $new_j, locked_j, dummy = 0$ ) has been received from  $G_j$  then
12:      foreach  $m \in [1 \dots n]$  do
13:        if ( $new_j[m] \neq \perp$ ) and ( $Vote_i[m] = \perp$ ) then
14:           $Vote_i[m] := new_j[m]$ ;  $new_i[m] := new_j[m]$ 
15:           $locked_i := locked_i \cup locked_j$ 
16:        endif
17:      enddo
18:    else
19:      if ( $G_j \notin locked_i$ ) then  $suspected_i := suspected_i \cup \{G_j\}$  endif
20:    endif
21:  enddo
22:  if ( $|suspected_i| > t$ ) then return (0) endif
23:  if ( $G_i \notin locked_i$ ) then
24:    if ( $r > |suspected_i|$ ) or ( $locked_i \neq \emptyset$ ) then  $locked_i := locked_i \cup \{G_i\}$  endif
25:  else
26:    if ( $|locked_i| > t$ ) then  $decide(Vote_i)$  endif
27:  endif
28: end_ round
29:  $decide(Vote_i)$ 
}
```

Procedure $decide(Vote)$ returns *decision* {

```
30: if ( $\exists m, 1 \leq m \leq n$ , s.t. ( $Vote[m] = \perp$ ) or ( $Vote[m] = 0$ )) then
31:   return (0)
32: else
33:   return (1)
34: end
}
```

Fig. 3. Pseudocode of a synchronous, early stopping Biased Consensus algorithm: code of process G_i .

message m to process G_j , but process G_j disregards m . Our solution assumes that all BC protocol messages have the same size.⁷ The motivation for having a full information protocol and uniform message size, will be explained in Section 4.

The changes we make to the algorithm of [28] do not affect the correctness of the algorithm. The difference is that we introduce the procedure $decide(\mathbf{Vote})$ (lines 30-34, Fig. 3), where \mathbf{Vote} is a vector of initially proposed values processes agree on. Basically, every process G_i stores information about the value initially proposed by some process G_j at $Vote_i[j]$. If the process G_i does not know the value that process G_j proposed, then $Vote_i[j] = \perp$. Roughly, a correct process does not learn the value proposed by process G_j , if G_j is faulty.⁸

We now give and prove the properties of our BC algorithm. We do not prove here that all processes that invoke $decide()$ function (lines 30-34) agree on the vector \mathbf{Vote} (this includes all correct processes, if $t < n/2$). Reader interested in this proof should refer to [28]. As discussed above, our algorithm inherently satisfies this property, given $t < n/2$. To summarize, our algorithm satisfies *Non Triviality*, *Validity* and *Biased Agreement* properties of BC. Furthermore, it satisfies the *Early Stopping* property:

- (Early Stopping) Every correct process decides in at most $\min(f + 2, t + 1)$ rounds.

Early Stopping property is inherited from the original algorithm. Consider *Non Triviality*. Assume that all processes are correct and all propose 1. In this case, all processes agree on vector $\mathbf{Vote} = 1^n$. Therefore $decide()$ returns 1 at every process. Consider now *Biased Agreement* and note that, if any process decides 1 it must have invoked $decide()$ and $\mathbf{Vote} = 1^n$. This implies that every correct process invokes $decide()$, evaluates the same vector \mathbf{Vote} that processes agreed on and, therefore, returns 1. Consider now *Validity*. If some process G_j proposed $vote_j = 0$ every process G_i that invokes $decide()$ (if any) has $Vote_i[j] = 0$ or $Vote_i[j] = \perp$, as processes agree on the vector \mathbf{Vote} and the coordinate j of \mathbf{Vote} is either \perp or $vote_j$. Therefore no process can decide 1. Note that *Validity* holds for any t .

3.4 Gracefully Degrading Fair Exchange

The impossibility of solving FE (deterministically) if half of the processes can be dishonest, motivates the introduction of the following variant of the problem.

Definition 3 (Gracefully Degrading Fair Exchange). *An algorithm solves gracefully degrading fair exchange (GDFE) if it satisfies the following properties:*

⁷ This could be implemented by meeting the maximal size of BC message by padding all message of BC algorithm to reach this size.

⁸ With different implementations of the $decide()$ procedure, the algorithm solves different problems. For example, if $decide(\mathbf{Vote})$ would return a minimum of all (non- \perp) coordinates, the algorithm would solve consensus, which is precisely the case in [28].

- The algorithm always satisfies the *Timeliness* and *Effectiveness* properties of fair exchange.
- If a majority of hosts are honest, then the algorithm also satisfies the *Fairness* property of fair exchange.
- Otherwise (if there is no honest majority), the algorithm satisfies *Fairness* with a probability p ($0 < p < 1$) such that the probability of unfairness $(1-p)$ can be made arbitrarily low.

4 A Gracefully Degrading Fair Exchange Algorithm

4.1 Description

Our GDFE algorithm is described in Figure 4. We assume that all processes involved in the algorithm know each other. The process with the lowest number is the initiator. We also assume a synchronous communication model [24] in the security subsystem. Basically, our algorithm can be viewed as an extension of the algorithm of Figure 2, i.e., our reduction of deterministic fair exchange to biased consensus. However, whereas the algorithm of Figure 2 is made of an *initialization* phase followed by an *agreement* (BC) phase, the algorithm of Figure 4 introduces a *fake* phase between these two phases. This is the key to graceful degradation, i.e., to minimizing the probability of unfairness in the case when $t \geq n/2$. Basically, we do not run the BC algorithm immediately after the exchange of items (i.e., unlike in Fig. 2), but at some randomly picked round. In the meantime the processes exchange *fake* messages and, if necessary, react to the behavior of hosts. If any process detects a host misbehavior, i.e., a message omission, it aborts the algorithm immediately (line 15) and does not participate in BC.⁹ It is important to notice that the underlying BC algorithm guarantees that no process decides 1 if some process does not participate in the algorithm (this missing process might have proposed 0). This is the way of penalizing any host that misbehaves in the first two phases of the algorithm.

The *Early Stopping* property of the underlying BC algorithm is essential for minimizing the probability for the adversary to violate fairness (as we discuss in the next subsection): in short, the early stopping BC algorithm we consider has two vulnerable rounds: if the adversary misses them, BC and the corresponding exchange terminate successfully. In addition, the introduction of *dummy* messages within the BC algorithm is necessary to solve the security requirement of our gracefully degrading fair exchange algorithm.

In our BC algorithm of Fig. 3, every process in every round sends exactly one message to every other process, but some (*dummy*) messages are tagged to be disregarded by the receiving process. This is necessary in order to make sure that the adversary has no means to distinguish the *fake* phase from the *agreement* phase (i.e., the BC algorithm), we make use of the same communication pattern

⁹ [25] uses a similar idea of choosing a random number of rounds and hiding it from the adversary to solve probabilistic non-repudiation (which is a special form of probabilistic fair exchange).

in both phases, i.e., the same distribution and sizes of the exchanged messages: Every process sends a fixed-size message to every other process in every round, both in the *fake* phase and in the BC algorithm. Messages in *fake* phase are, therefore, padded before sending, to the size of BC message. Hence, the adversary is not able to determine when BC starts, neither by observing when security modules send and receive messages, nor by observing the size of these messages.

```

GDFairExchange(myitem, description, source, destination) returns item {
01: if  $\langle G_i \text{ is initiator} \rangle$  then % initialization phase - round 0
02:    $\langle$ pick a random number  $k$  according to a given distribution $\rangle$ 
03:   foreach  $G_j \neq \text{destination}$  do send  $(\perp, k)$  to  $G_j$  enddo
04:   send (myitem,  $k$ ) to destination
05: else
06:   send (myitem,  $\perp$ ) to destination
07: endif
08: if  $((\text{item}, *) \text{ has been received from } \text{source})$  and
    $(\text{item matches } \text{description})$  and  $((*, k) \text{ has been received from } \text{initiator})$  then
09:    $\text{vote}_i := 1$ ;  $k_i := k$ ;  $\text{item}_i := \text{item}$ 
10: else
11:   return  $(\perp)$ 
12: endif
13: for  $\text{round} := 1, 2, \dots, k_i$  do % fake phase -  $k$  rounds
14:   send  $\langle$ padded  $\text{vote}_i$  $\rangle$  to all
15:   if not $((\text{vote}) \text{ has been received from all processes})$  then return  $(\perp)$  endif
16: enddo
17:  $\text{vote}_i := \text{BCpropose}(\text{vote}_i)$  % agreement phase - Biased Consensus
18: if  $(\text{vote}_i = 1)$  then return (item $i$ ) else return  $(\perp)$  endif
}

```

Fig. 4. Pseudocode of the Gracefully Degrading Fair Exchange algorithm: code of process G_i .

4.2 Correctness of GDFE Algorithm

Theorem 4. *The algorithm of Figure 4 solves gracefully degrading fair exchange.*

Proof. The *Timeliness* property is guaranteed by the fact that we consider a synchronous system and the *Termination* property of BC.

Consider *Effectiveness* and assume that all participating hosts are honest and all items match their descriptions. All security modules will enter and exit the *fake* phase having $\text{vote} = 1$, so all security modules will *BCpropose* 1. By the *Non Triviality* property of BC every module returns 1 and subsequently returns the expected item to its host.

Now we consider *Fairness*. It is important here to recall that the security modules are tamper-proof and no information leaks from them apart from what

is explicitly released through their interface. We first prove a preliminary lemma. For convenience, if the security module returns \perp to its host, we say that security module aborts the GDFE algorithm.

Lemma 1. *If the first round in which some security module G_j aborts the GDFE algorithm is round i ($0 \leq i < k$), then at the end of round $i + 1$ every security module has aborted the GDFE algorithm.*

Proof. Because G_j has aborted the GDFE algorithm at the end of round i , no security module will receive G_j 's *vote* in round $i + 1$. From line 15, it can be seen that every security module will abort the algorithm at latest at the end of round $i + 1$ (some modules might have aborted the algorithm in round i , like G_j). \square

Consider the case in which the first misbehavior of some of the dishonest hosts occurs in the round i where $0 \leq i < k$ (misbehavior in round 0 includes the initiator's misbehavior or some dishonest host sending the wrong item). According to Lemma 1, by the end of the round $i + 1 \leq k$, all security modules will abort the algorithm, so *Fairness* is preserved.

Note that Lemma 1 does not hold for the k -th round. Some dishonest hosts can cut the channels for the first time in that round in such way that some security modules receive all messages and some do not. Hence some modules will *BCpropose* 1 and others will abort the algorithm at the end of round k and will not participate in BC. Because the modules that invoked consensus cannot distinguish this run from the run in which some faulty module proposed 0 and failed immediately in such way that it did not send or receive any message, all security modules that had invoked BC will return 0. At the end, none of the hosts gets the item.

The last possibility is that the first misbehavior occurs during the execution of the BC. This means that every security module has proposed 1 to BC. If there is a majority of honest hosts, the *Biased Agreement* property of BC guarantees *Fairness*. Indeed, *Fairness* can be violated only if some security module returns the expected item to its host, while some correct security module returns \perp to its honest host. From line 18, it is obvious that this would be possible only if some security module returns 1 from BC, while some correct security module returns 0 which contradicts the *Biased Agreement* property.

If the adversary controls half or more of the hosts *Fairness* could be violated if, and only if, the adversary cuts one of the first two rounds of BC. However, this could occur only if the adversary successfully guesses in which round BC starts. Indeed, because our BC algorithm is *early stopping*, in order to succeed, the adversary must cut one of the first two rounds of BC and this has to be its first misbehavior in a particular algorithm run. In the following, we prove that if this case occurs, i.e., if there is no honest majority, probability of unfairness can be made arbitrarily low by choosing an appropriate distribution of the random number of *fake* rounds.

The number k of rounds in the second phase of the algorithm is chosen randomly by the initiator of the exchange according to a given distribution $(\beta_0, \beta_1, \dots)$ i.e., $\Pr(k = i) = \beta_i$. We assume this distribution to be public. The

adversary performs the attack in a given round by dropping a certain subset of messages sent to, or received by, the hosts it controls, i.e., by cutting the channels. When the adversary cuts channels at more than $n/2$ hosts in the same round, we say that he *cuts the round*. Since the adversary does not know in which round BC starts, the best attack consists in choosing a value i according to the distribution $(\beta_0, \beta_1, \dots)$, starting from which adversary cuts all the rounds until the end of the exchange. Cutting messages at less than $n/2$ hosts, or cutting non-consecutive rounds, cannot improve the probability of success of the adversary.

We define the *probability of unfairness* $\Gamma_{(\beta_0, \beta_1, \dots)}$ as the maximum probability that an adversary succeeds, given the distribution $(\beta_0, \beta_1, \dots)$, and the *average complexity* in terms of number of *fake* rounds as $\Lambda_{(\beta_0, \beta_1, \dots)} = \sum_{i \geq 1} i \beta_i$.

Lemma 2. *Let $(\beta_0, \beta_1, \dots)$ denote the probability distribution of the value k . The probability of unfairness (for the algorithm of Figure 4) is*

$$\Gamma_{(\beta_0, \beta_1, \dots)} = \max_{i \geq 0} (\beta_i + \beta_{i+1}).$$

Proof. Let γ_i be the probability that the attack succeeds if it starts at round i ($i > 0$). We already know that $\gamma_{i \leq k} = 0$ and that $\gamma_{i > k+2} = 0$. We have therefore:

$$\gamma_1 = \beta_0, \gamma_2 = \beta_0 + \beta_1, \gamma_3 = \beta_1 + \beta_2, \dots, \gamma_i = \beta_{i-2} + \beta_{i-1}, \dots$$

According to the probability distribution $(\beta_0, \beta_1, \dots)$, the maximum probability of unfairness $\Gamma_{(\beta_0, \beta_1, \dots)}$ is therefore $\Gamma_{(\beta_0, \beta_1, \dots)} = \max_{i > 0} (\gamma_i) = \max_{i \geq 2} (\beta_0, \beta_{i-2} + \beta_{i-1}) = \max_{i \geq 0} (\beta_i + \beta_{i+1})$. \square

We define the probability distribution that we call *bi-uniform*, as well as the *optimal* probability distribution for the algorithm of Figure 4.

Definition 4. *We say that $(\beta_0, \beta_1, \dots)$ is a bi-uniform probability distribution of parameter t on the interval $[0, \kappa]$ if $\forall i \geq 0, \beta_i + \beta_{i+1} = \frac{1}{\lfloor \frac{\kappa+1}{2} \rfloor}$ and $\beta_1 = t$ if κ is odd, and $\beta_1 = 0$ if κ is even.*

Definition 5. *We say that a probability distribution $(\beta_0, \beta_1, \dots)$ is optimal (for the algorithm of Figure 4) if there is no other probability distribution $(\beta'_0, \beta'_1, \dots)$ such that $\exists \Gamma > 0, \forall i \geq 0, \beta_i + \beta_{i+1} \leq \Gamma, \beta'_i + \beta'_{i+1} \leq \Gamma$ and $\Lambda_{(\beta'_0, \beta'_1, \dots)} < \Lambda_{(\beta_0, \beta_1, \dots)}$.*

The following lemma states our optimality result in terms of probability of unfairness.

Lemma 3. *The optimal probability distribution (for the algorithm of Figure 4) is the bi-uniform probability distribution of parameter 0. Moreover, if the distribution is defined on $[0, \kappa]$ with κ even, the probability of unfairness is $\Gamma_{\text{bi-uniform}} = \frac{2}{\kappa+2}$ and the average complexity, in terms of the number of fake rounds, is $\Lambda_{\text{bi-uniform}} = \frac{\kappa}{2}$.¹⁰*

Due to the lack of space, we omit here the formal proof of Lemma 3. The proof can be found in the full version of the paper [2]. \square

¹⁰ It can be shown that our GDFE algorithm satisfies $\Gamma_{\text{bi-uniform}} / \Xi_{\text{bi-uniform}} \approx 2 / \Lambda_{\text{bi-uniform}}$, where Ξ is the probability that all processes successfully terminate

5 Concluding Remarks

It was shown in [16] that deterministic two-party fair exchange is impossible without a trusted third party. Published results on multi-party protocols focus on reducing the necessary trust in the third party [7, 20], or on contract signing [8, 21], a special form of fair exchange. In [22] it was shown that, in a synchronous system with computational security, a majority of honest processes can simulate a centralized trusted third party (and hence solve fair exchange) using cryptography. In a somewhat stronger model, [18] gives the solution for secure multi-party computation in a synchronous system with unconditional security, that also assumes a majority of honest processes. The use of security modules in fair exchange was already explored in the two-party context: in particular, [34] employs smart cards as security modules to solve two-party fair exchange in an optimistic way, whereas [3] describes a probabilistic solution to two-party fair exchange.¹¹ Idea of using a distributed trusted third party in solving two-party fair exchange was exploited in [4].

Recent works [19, 23] have solved various forms of secure multi-party computation (SMPC) for any number of dishonest parties ($t < n$). As fair exchange is usually considered as a special case of a SMPC, it can be tempting to conclude that these results also apply to fair exchange. However, in the relaxed definitions of SMPC considered in [23] and, implicitly in [19], *fairness* is not always required, as discussed in [23]. We consider in this paper contexts where *fairness* is mandatory. In this sense, results shown in this paper are rather complementary to those that establish the possibility of solving SMPC for any $t < n$.

Acknowledgments

We are very grateful to Bastian Pochon for his useful comments on the specification of biased consensus. Gildas Avoine is supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

the algorithm; rephrasing the result of [33] in our context would mean that the optimal for a randomized biased consensus is $\Gamma/\varepsilon \geq 1/\Lambda$ (instead of $2/\Lambda$). We claim that the factor 2 in our case is due to the fact that we ensure biased agreement with a correct majority.

¹¹ Two-party fair exchange is clearly simpler than the multi-party fair exchange. Informally, a deterministic solution with honest majority in the two-party case, i.e., when both (all) participating hosts are honest, is possible in a single communication round. On the other hand, two rounds of communication are minimum for the deterministic solution to biased consensus (and, therefore, for fair exchange) with the honest majority.

References

1. N. Asokan, M. Schunter, and M. Waidner. Optimistic protocols for fair exchange. In T. Matsumoto, editor, *4th ACM Conference on Computer and Communications Security*, pages 8–17, Zurich, Switzerland, Apr. 1997. ACM Press.
2. G. Avoine, F. Gärtner, R. Guerraoui, and M. Vukolić. Gracefully degrading fair exchange with security modules. Technical Report IC/2004/26, EPFL, Switzerland, 2004.
3. G. Avoine and S. Vaudenay. Fair exchange with guardian angels. In *The 4th International Workshop on Information Security Applications – WISA 2003*, Lecture Notes in Computer Science, Jeju Island, Korea, August 2003. Springer-Verlag.
4. G. Avoine and S. Vaudenay. Optimistic fair exchange based on publicly verifiable secret sharing. In *Proceedings of the 9th Australasian Conference on Information Security and Privacy*, Lecture Notes in Computer Science, pages 74–85, Sydney, Australia, 2004. Springer-Verlag.
5. M. Backes, B. Pfizmann, M. Steiner, and M. Waidner. Polynomial fairness and liveness. In *15th IEEE Computer Security Foundations Workshop – CSFW*, pages 160–174, Cape Breton, Nova Scotia, Canada, June 2002. IEEE, IEEE Computer Society.
6. F. Bao, R. Deng, and W. Mao. Efficient and practical fair exchange protocols with off-line TTP. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 77–85, Oakland, California, USA, May 1998. IEEE Computer Society Press.
7. F. Bao, R. Deng, K. Q. Nguyen, and V. Varadharajan. Multi-party fair exchange with an off-line trusted neutral party. In *Proceedings of the 10th International Workshop on Database & Expert Systems Applications*, pages 858–862, Florence, Italy, 1–3 Sept. 1999. IEEE Computer Society Press.
8. B. Baum-Waidner and M. Waidner. Round-optimal and abuse-free multi-party contract signing. In *27th International Colloquium on Automata, Languages and Programming (ICALP '2000)*, volume 1853 of *Lecture Notes in Computer Science*, pages 524–535, Geneva, Switzerland, July 2000. Springer-Verlag.
9. P. Bernstein, V. Hadzilacos, and N. Goodman. *Concurrency Control and Recovery in Database Systems*. Addison-Wesley, Reading, MA, 1987.
10. D. Boneh and M. Naor. Timed commitments. In M. Bellare, editor, *Advances in Cryptology – CRYPTO 2000*, volume 1880 of *Lecture Notes in Computer Science*, pages 236–254, Santa Barbara, California, USA, August 2000. IACR, Springer-Verlag.
11. C. Boyd and E. Foo. Off-line fair payment protocols using convertible signatures. In K. Ohta and D. Pei, editors, *Advances in Cryptology – ASIACRYPT'98*, volume 1514 of *Lecture Notes in Computer Science*, pages 271–285, Beijing, China, October 1998. IACR, Springer-Verlag.
12. C. Boyd and P. Kearney. Exploring fair exchange protocols using specification animation. In J. Pieprzyk, E. Okamoto, and J. Seberry, editors, *Proceedings of ISW 2000*, volume 1975 of *Lecture Notes in Computer Science*, pages 209–223, Wollongong, Australia, December 2000. Springer-Verlag.
13. H. Bürk and A. Pfizmann. Value exchange systems enabling security and unobservability. *Computers & Security*, 9(8):715–721, 1990.
14. L. Chen. Efficient fair exchange with verifiable confirmation of signatures. In K. Ohta and D. Pei, editors, *Advances in Cryptology – ASIACRYPT'98*, volume 1514 of *Lecture Notes in Computer Science*, pages 286–299, Beijing, China, October 1998. IACR, Springer-Verlag.

15. J. G. Dyer, M. Lindemann, R. Perez, R. Sailer, L. van Doorn, S. W. Smith, and S. Weingart. Building the IBM 4758 secure coprocessor. *IEEE Computer*, 34(10):57–66, Oct. 2001.
16. S. Even and Y. Yacobi. Relations among public key signature systems. Technical Report 175, Computer Science Department, Technion, Haifa, Israel, 1980.
17. M. J. Fischer, N. A. Lynch, and M. S. Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2):374–382, Apr. 1985.
18. M. Fitzi, N. Gisin, U. M. Maurer, and O. von Rotz. Unconditional byzantine agreement and multi-party computation secure against dishonest minorities from scratch. In *Proceedings of the 21st annual symposium on the Theory and Applications of Cryptographic Techniques*, pages 482–501. Springer-Verlag, 2002.
19. M. Fitzi, D. Gottesman, M. Hirt, T. Holenstein, and A. Smith. Detectable byzantine agreement secure against faulty majorities. In *Proceedings of the twenty-first annual symposium on Principles of distributed computing*, pages 118–126. ACM Press, 2002.
20. M. K. Franklin and G. Tsudik. Secure group barter: Multi-party fair exchange with semi-trusted neutral parties. In *Financial Cryptography – FC ’98*, volume 1465 of *Lecture Notes in Computer Science*, pages 90–102, Anguilla, British West Indies, Feb. 1998. Springer-Verlag.
21. J. A. Garay and P. MacKenzie. Abuse-free multi-party contract signing. In *Distributed Computing – DISC ’99*, volume 1693 of *Lecture Notes in Computer Science*, pages 151–165, Bratislava, Slovak Rep., 27–29 Sept. 1999. Springer-Verlag.
22. O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game — a completeness theorem for protocols with honest majority. In *Proceedings of the 19th ACM Symposium on the Theory of Computing (STOC)*, pages 218–229, 1987.
23. S. Goldwasser and Y. Lindell. Secure computation without agreement. In *Proceedings of the 16th International Conference on Distributed Computing*, pages 17–32. Springer-Verlag, 2002.
24. N. Lynch. *Distributed Algorithms*. Morgan Kaufmann, San Mateo, CA, 1996.
25. O. Markowitch and Y. Roggeman. Probabilistic non-repudiation without trusted third party. In *Proceedings of the 2nd Workshop on Security in Communication Networks*, 1999.
26. J. McLean. A general theory of composition for a class of “possibilistic” properties. *IEEE Transactions on Software Engineering*, 22(1):53–67, Jan. 1996. Special Section—Best Papers of the IEEE Symposium on Security and Privacy 1994.
27. H. Pagnia, H. Vogt, and F. C. Gärtner. Fair exchange. *The Computer Journal*, 46(1), 2003.
28. P. R. Parvédy and M. Raynal. Optimal early stopping uniform consensus in synchronous systems with process omission failures. In *Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures*, pages 302–310. ACM Press, 2004.
29. K. J. Perry and S. Toueg. Distributed agreement in the presence of processor and communication faults. *IEEE Transactions on Software Engineering*, 12(3):477–482, Mar. 1986.
30. D. Skeen. Non-blocking commit protocols. In *Proc. ACM SIGMOD Conf.*, page 133, Ann Arbor, MI, Apr.-May 1981.
31. T. Tedrick. Fair exchange of secrets. In G. R. Blakley and D. Chaum, editors, *Advances in Cryptology – CRYPTO’84*, volume 196 of *Lecture Notes in Computer Science*, pages 434–438, Santa Barbara, California, USA, August 1985. IACR, Springer-Verlag.

32. Trusted Computing Group. Trusted computing group homepage. Internet: <https://www.trustedcomputinggroup.org/>, 2003.
33. G. Varghese and N. A. Lynch. A tradeoff between safety and liveness for randomized coordinated attack. *Information and Computation*, 128(1):57–71, 1996.
34. H. Vogt, F. C. Gärtner, and H. Pagnia. Supporting fair exchange in mobile environments. *ACM/Kluwer Journal on Mobile Networks and Applications (MONET)*, 8(2), Apr. 2003.