J. L. Ferrer-Gomilla, J. A. Onieva, M. Payeras, and J. Lopez, "Certified electronic mail: Properties revisited", *Computers & Security*, vol. 29, pp. 167 - 179, 2010.
http://doi.org/DOI:10.1016/j.cose.2009.06.009
NICS Lab. Publications: https://www.nics.uma.es/publications

CERTIFIED ELECTRONIC MAIL: PROPERTIES REVISITED

Abstract

Certified electronic mail is an added value to traditional electronic mail. In the definition of this service some differences arise: a message in exchange for a reception proof, a message and a non repudiation of origin token in exchange for a reception proof, etc. It greatly depends on whether we want to emulate the courier service or improve the service in the electronic world. If the definition of the service seems conflictive, the definition of the properties and requirements of a good certified electronic mail protocol is even more difficult. The more consensuated features are the need of a fair exchange and the existence of a trusted third party (TTP). Each author chooses the properties that considers the most important, and many times the list are conditioned by the proposal. Which kind of TTP must be used? Must it be verifiable, transparent and/or stateless? Which features must the communication channel fulfill? Which temporal requirements must be established? What kind of fairness is desired? What efficiency level is required? Are confidentiality or transferability of the proofs compulsory properties? In this paper we collect the definitions, properties and requirements related with certified electronic mail. The aim of the paper is to create a clearer situation and analyze how some properties cannot be achieved simultaneously. Each protocol designer will have to decide which properties are the most important in the environment in where the service is to be deployed.

1. INTRODUCTION

The consensus amongst the security and cryptographic applications development community is that the society demands value-added services over the existing electronic services. In the case of the electronic mail there is a trivial extension: *Certified Electronic Mail (CEM)*. This consensus does not exist with respect to the definition of this service nor to the properties it needs to fulfil in order to be considered a useful digital service. There are multiple different opinions and nowadays legislators, enterprises (implementing certified email applications) and the research community itself does not reach an agreement. More surprising may be the fact that the research community has not reached a common view either.

There is a lack of international specific regulations for certified electronic mail. An EU directive has been issued on community postal services (EU directive 97/67/CE), however its scope is not considered to include electronic mail. Some specific domestic legislation covers the legal validity of certified electronic mail. Some aspects that can be addressed by this legislation can be the evidence types and its content, the involved actors and its responsibilities.

Among the enterprises implementing certified email applications, there is a consensus of which is the essential aspect of certified email: the provision of evidence of actions relating with the handling on messages [ETSI]. Almost everybody agree in the fact that the products must support evidence of message origin authentication and evidence of submission. Other kinds of evidence are not as consensual, as the evidence of notification to the recipient of the availability of a dropped message or the evidence of delivery/download. Other open questions are the identity of the part that generates the evidences or how the evidences are carried.

Perhaps the best way to deal with CEM service is to firstly check up on the paper world. Certified postal mail can be considered as a service in which the post office (or postman representing the post office) does not deliver a mail unless the recipient signs a receipt. This exchange is fair, and its fairness is ensured by physical presence. Its electronic version needs to provide, as far as possible, the same properties this service provides in the paper world but, obviously, avoiding the physical presence requirement.

We can extract from all the related literature [AG02, Ateniese01, FG02, DGLW96, OZL04, PFR05, WBZ04, Zhou01, ZOL05] that there is the agreement that certified electronic mailing should be a fair exchange of items (consensus, which is at the same time discussable as we will see later on). There is not such an agreement with respect to which items are to be exchanged. Let us begin with this definition:

Definition 1. The certified electronic mail service can be defined as the exchange of a message plus an evidence of origin for an evidence of receipt.

Indeed the latter definition does not differ much from the definition provided for non-repudiation protocols. We could, thus, recall the following definitions related to non-repudiation protocols:

Definition 2 (Non-repudiation of Origin NRO). A protocol provides a non-repudiation of origin service if, and only if, it generates evidence of origin which will allow the recipient to demonstrate to an arbiter whether the originator was or not the message's author.

Definition 3 (Non-repudiation of Receipt NRR). A protocol provides a non-repudiation of receipt service if, and only if, it generates evidence of receipt which will allow the originator to demonstrate to an arbiter whether the recipient received the message.

Some authors [ZG96] do not agree with **definition 1**. Obviously the definition conditions the possible solutions which can be implemented for a CEM service. These authors allege that the equivalent postal service does not require the originator creating evidence of origin and thus, they prefer this definition:

Definition 4. The certified electronic mail service can be defined as the exchange of a message for an evidence of receipt.

Therefore, the evidence of origin is an optional property (hence the protocol doesn't need to fulfil **definition 2**). We agree that the equivalent postal service does not require evidence of origin to be generated, but we also admit that in general, the originator sends evidence of origin (e.g. its hand-written signature in the message).

Additionally we could discuss the evidence of receipt concept. Generally the postal services force the recipient to sign a receipt token before delivering the envelope which contains the certified message; but with this signature the recipient only recognizes that it received an envelope which, in turn, can be empty (intentionally or not)[1]. Hence there is a difference with respect to the digital evidence of receipt (linked to the message and not to the envelope). This leads to a definition closer to the postal certified service:

Definition 5. The certified electronic mail service can be defined as the exchange of a digital envelope (which supposedly contains a message) for an acknowledgement of receipt (not linked with the supposed received message).

From this definition we can conclude that a CEM protocol does not need to fulfil **definitions 2 (non-repudiation of origin)** and **3 (non-repudiation of receipt)**. It is arguable whether that service is useful or not, but it is out of discussion that this definition is more similar to the postal service. From this debate we can extract the following combinations:

1. Exchange of message and NRO for NRR linked to the message.
2. Exchange of message and NRO for acknowledgement of receipt.
3. Exchange of message for NRR linked to the message.
4. Exchange of message for acknowledgment of receipt.

[1] We need to distinguish between certified mail postal service and other certified delivery services; e.g. *burofax*. In some countries like Spain this service provides the originator a signed copy of the received message.

5. Exchange of envelope and, if possible, NRO for NRR, if possible, linked to the message.
6. Exchange of envelope and, if possible, NRO for acknowledgement of receipt.
7. Exchange of envelope for NRR, if possible, linked to the message.
8. Exchange of envelope for acknowledgement of receipt.

Therefore, bearing in mind only the elements to be exchanged in the certified electronic mail service, not considering the properties to be satisfied, we already find eight different services.
We would like to highlight that although all combinations are possible, perhaps not all of them would make sense in the digital world. After a first selection process we will have the practical definitions, and successively, we will try to work out the best definition. Let us put forward that the first definition (**definition 1**) seems to be more practical albeit others could be of interest in certain scenarios.

Another related service in the paper world is the notification service. In fact, sometimes there is no difference between a notification and a certified mail (from the law point of view there is). Nevertheless there is a special type of notification. Some notifications force the addressee to provide evidence of receipt that needs to be signed over a carbon copy. This specific type of notification is best defined by **definition 1**. The main difference to the electronic service is that the delivery agent sets the document as refused if the user declines to accept the notification (for legal purposes). Unfortunately this cannot be done in the electronic world since it is not possible to distinguish between the following different cases:

- The recipient declines to accept the message.
- There is a channel failure which prevents the user from getting the message.

In order to mitigate this issue, we can try to reduce the possibility of the originator producing a selective receipt based on the message content or authorship. The first one is ensured by fairness property (see next section). The latter can be obtained by hiding the originator's identity and making use of an anonymous channel as suggested by [KM01]. Hence this protocol provides users with a service as close as possible to the one provided by the post in which the postman only delivers the mail (and reveals at that moment the sender's identity if it is written on the envelope) when he has obtained the signed receipt for it. Later Gonzalez-Deleito introduced a variant which allowed reaching the same property without putting full trust in the TTP[2] (but unfortunately, as the author states, the protocol does not guarantee the delivery of a non-repudiation of origin evidence to the recipient [Nicolas05]).

[2] Trusted Third Party

Note that this does not prevent the certified electronic mail from having two possible cases either. The receiver can still reject the reception of the email, but he will do so without any information on what and from whom he is refusing to receive the message. This renders refusal to receive a message almost useless. Nevertheless, although this presumably accounts for a service with less selective receipts than in the paper world, it still suffers from legal ambiguity due to possible channel failures.

Let us now introduce all the existing properties which will allow us to classify the different CEM protocols and point out the kind of certified electronic mail service they provide.
Table 1 summarizes these properties. They will be defined and described in the paper, following this classification. Furthermore, we will also stress whether these properties are mandatory, optional, desirable, etc.

Category	**Properties (Definition Number)**	**Comments**
Non Repudiation	Non repudiation of origin (2) Non repudiation of receipt (3)	Both properties are desirable. Depending on the definition of certified e-mail they are both mandatory
Fairness	Fairness / Strong Fairness (6) Weak Fairness (7) Light Fairness (8) True Fairness (9) Probabilistic Fairness (10)	Fairness is mandatory, so one of these properties must be compulsory. Weak Fairness is enough, although strong fairness is desirable Probabilistic fairness is not desirable
TTP	In-line TTP (11) On-line TTP (12) Off-line TTP (13) Transparent TTP (14) Verifiable TTP (15)	Off-line TTP is desired, but the involvement of the TTP depends of the application. Transparent and Verifiable TTP are desired, but only one of them can be achieved because they are incompatible
Communication Channel	Operational Channel (16) Unreliable Channel (17) Resilient Channel (18)	A resilient channel between the parts is desirable, better if it can be unreliable. An unreliable channel between the parts and the TTP is desirable.
Temporality	Timeliness (19) Synchronous Timeliness (20) Asynchronous Timeliness (21)	Asynchronous Timeliness is desirable.
State Storage	Strong Stateless TTP (22) Weak Stateless TTP (23) Strong Stateful TTP (24)	Strong Stateless is the most desirable.

	Weak Stateful TTP (25)	
Others	Efficiency (26) Evidence Transferability (27) Confidentiality (28)	These properties are desirable.

Tabla 1. Classification of Properties

Fairness properties are described in section 2. The properties related with the TTP are included in sections 3 and 4. Section 5 includes the properties related with communication channels and timeliness. State storage is classified in section 5. Finally, the remaining properties are described in sections 7, 8 and 9.

2. FAIRNESS

This is a core property for CEM applications. A fair protocol ensures that no party will be in an advantageous position with respect to another one when participating in an exchange of items (whether this exchange consists in a message and NRO with NRR, etc.). But even this property can have different approximations. Maybe the most appropriate definition is:

Definition 6 (Fairness). A CEM protocol fulfils the fairness property if, and only if, by the end of the execution

a) both, sender and recipient, receive the expected items (or will receive them in a finite amount of time); or
b) none of them receives what it expects.

Some authors name this property (as previously defined) **strong fairness** [Aso98]. But even this property allows alternative definitions [Aso98]:

Definition 7 (Weak fairness). A CEM protocol is said to be weakly fair if, and only if, by the end of the execution

a) both, sender and recipient, receive the expected items; or
b) if one entity receives the expected item and another entity does not, the latter can get evidence of this situation.

This definition allows obtained evidence to be invalidated. Hence fairness is not weak but conditionally verifiable. That is, all entities need to provide evidence to the arbiter if there is a dispute resolution process.

Definition 8 (Light fairness). This definition of fairness can be applied to protocols with more than one recipient. It appears in [OLZ09]. A multi-party CEM protocol is said to provide light fairness if, and only if:

a) originator and recipient obtain NRR and NRO respectively
b) none of them obtains any evidence

Note that this property only ensures fair exchange of evidence, allowing the recipient to get the message without providing evidence of receipt (but, at the same time, without obtaining evidence of origin).

A priori these two definitions cannot be considered as valid for CEM applications, because the exchange is not fair if, by the end of the protocol execution, one party possesses an item which places him in a more advantageous position. Nevertheless in some applications with additional restrictions these definitions of fairness could be sufficient.

Some existing definitions are of limited accuracy. For instance, Kremer et al. [KMZ02] proposes the following definition in the context of certified electronic mail:

Definition 9 (True fairness). A CEM protocol is truly fair if, and only if, it fulfils strong fairness, and if the exchange has been successful, generated evidence is independent from how the protocol was executed.

In other words, these authors consider that a protocol is truly fair if the evidence generated by a TTP cannot be distinguished from evidence generated by sender and recipient when the TTP does not participate. Therefore true fairness ranks equally with another property that protocols can (or cannot) fulfil: the TTP being **transparent**. In the following sections we will argue that this property is optional and not compatible with other properties that could be even more important (at least in the certified electronic mail context).

On the other hand, from a scientific point of view we also have the following definition of fairness:

Definition 10 (Probabilistic fairness). A CEM protocol is probably fair with a probability ε if, and only if, it fulfils fairness as defined in **definition 6**, or the probability of a cheating party gaining valuable information of its expected items, while the other party gains nothing, is ≤ ε.

We consider probabilistic fairness [MR99] to be less practical because, although it presents an interesting approach from the scientific point of view, it is practically not acceptable by users and arbiters. That is, even if ε is very small, most solutions require a deterministic fairness property.

3. TRUSTED THIRD PARTIES

In this paper we will not take into account those approaches that try to remove the participation of a TTP since none of them is practically feasible. Some of these solutions are not fair and make assumptions like equal computational power among the protocol participants. Thus assuming the TTP will always be part (to a greater or lesser extent) of the CEM application, we could classify its participation in the following way:

Definition 11 (In-line TTP). A TTP is said to be in-line if it acts as an intermediary in all the exchanges among users during the protocol execution.

Definition 12 (On-line TTP). A TTP is said to be on-line if it acts in all protocol executions.

Definition 13 (Off-line TTP). A TTP is said to be off-line if it does not participate actively in the protocol; i.e. it will only be invoked when there is an exception in the transaction.

Schemes that use the latter definition are also named **optimistic**, and it is, nowadays, the most used paradigm when designing CEM protocols (but also non-repudiation, digital contract signing protocols, etc.). This is due to disadvantages the in-line and on-line solutions introduce:

- the TTP could become a bottleneck as it intervenes in every exchange, and
- the service will be more expensive if the TTP charges per intervention.

Obviously these disadvantages are worse in the case of in-line solutions. Both issues are related. Strictly, the first hindrance is not completely true. Currently multiple companies use the Internet as a platform to provide their services (search engines as Google, auction sites as eBay, on-line

retailers as Amazon), and their servers accept millions of requests per day without becoming a bottleneck. However, it is true that this will require a more expensive infrastructure to put in place and maintain.

Although in the optimistic schemes the TTP acts only in case of misbehaviour or channel failure, due to its continuous presence it exercises a coercive role against cheating entities. In any case, the TTP needs to maintain an appropriate infrastructure in order to answer the incoming requests. So even if the costs are calculated per intervention, this TTP has a maintenance cost as well (though less expensive than in-line and on-line schemes). It is subject to study whether the cost difference is large enough to influence the final selection (and thus whether it is a determinant argument).

Optimistic solutions are more elegant and efficient and research efforts need to be applied in this type of schemes. But the existing Internet applications show a different model. Several countries and companies provide a certified electronic mail service. All the analyzed schemes use an in-line TTP (the company or institution providing the service) and some of them do not fulfil properties which are very relevant for CEM applications (e.g. [PEC,PIT]). This makes us wonder whether research should be conducted towards on-line and in-line applications fulfilling all the expected and intended CEM properties (e.g. strong fairness) such that companies could provide services with validated schemes.

In this sense, and searching for a trade-off, it might be that on-line solutions present the best market-based model in order to be successful on the Internet. Other researchers agree on this statement [Oppliger04]. Users may need to trust the TTP (see next section) in every exchange, but this, though being an obstacle from the research point of view, could, on the other hand, facilitate the use and the comprehension of the users of the infrastructure they are using. At the same time this kind of solutions, as we explained before, are less sensitive to turn the TTP into a bottleneck. Besides the cost of deployment will be less than that of in-line solutions.

Therefore, real Internet applications and conducted research must match requirements and solutions in order to come up with an optimal type of TTP.

4. TRANSPARENCY AND VERIFIABILITY OF THE TTP

When TTP participates in an off-line manner, other properties are of interest in the application: transparency and verifiability. Let us analyze the first one:

Definition 14 (Transparent TTP). It also appears in the literature as *invisible* TTP [Micali97]. A TTP is said to be transparent if at the end of a successful protocol execution collected evidence makes impossible to differentiate whether the TTP was involved or not.

Hence, the resulting evidence will be the same as the one obtained in the case the TTP is not involved. Some authors consider this property as especially important for optimistic schemes. In these schemes this property helps on the privacy of users with respect to the use or not of a TTP during the protocol run. When the TTP gets involved in a protocol execution either a party misbehaved or there was a channel failure. As previously stated it is theoretically impossible to distinguish between both cases, thus raising the suspicion in all cases about whether the entities were misbehaving. Therefore, this is the reason for which some authors regard this property as essential, particularly in electronic commerce applications as it avoids bad (and not demonstrable) reputation. Furthermore, this could be important in practical cases, in which an institution does not wish to change the existing processes to accommodate the new signatures or affidavits generated by TTPs.

On the other hand, there can be scenarios in which the TTP could not be completely trusted [LNJ00]. All entities (humans, machines and computer programs) are subject to errors. Hence, the TTP could benefit or damage one or several honest users. This problem is not new to the digital world. In our daily life there exist trusted third parties: arbiters, notaries, police, delivery agents, judges, etc. And all of them may misbehave: judges taking a bribe and/or pronouncing unfair sentences, a policeman who overstep the bounds, etc. In our conventional world we have created procedures to avoid (or at least to restrict) this misbehaviour: reasoned signed sentences that can be appealed, signed accusations that can be appealed by the affected entity, etc. Even more, it is important to generate evidence, not only because of misbehaviour, but also because of possible unintentional errors carried out by third parties. In our digital world error probability can be reduced but never dismissed and, at the same time, TTPs can also be corrupted. Therefore the following property needs to be taken seriously into account:

Definition 15 (Verifiable TTP). A TTP is said to be verifiable when involved in a certified electronic mail protocol if it generates evidence which allows demonstrating its participation.

That is, the service provided by the TTP is verifiable. In [PFR05], a TTP is verifiable if the services it provides are verifiable, and the verifiability of the services is *on-line*. In this paper, the authors distinguish two different verifiability options. Definitions are reproduced as they originally appear in [PFR05]:

- *On-line Verifiability*: a service is on-line verifiable when a user can immediately know whether the TTP misbehaved by checking the evidences received from the TTP. In case of problems, the user can start a dispute to correct the situation.
- *Off-line Verifiability*: the verifiability of a security service is off-line when the evidences received from a TTP are not enough to know if it has been provided properly or not. But if a dispute arises between the parties involved in the protocol, then the evidences can be used to prove whether the TTP misbehaved.

In our definition, a TTP is verifiable if the verifiability of the services it provides is *off-line*. As authors state, in order to provide on-line (real-time) verifiability, strong assumptions need to be made (e.g. deployment of a public directory which is not controlled by the TTP).

That is, it is not possible to completely avoid TTP's misbehaviour or unintentional errors, but it is important to allow honest users, at least, to demonstrate this fact (thus avoiding an unfair termination of the protocol).

Once we introduced both properties, we will demonstrate that they are not compatible:

Motto 1. If a TTP is transparent (**definition 14**) it cannot be verifiable (**definition 15**).

Proof. Definition 14 states that TTP's generated evidence cannot be distinguished from generated evidence when it does not directly participate. Hence, it is not possible to demonstrate whether it was involved in the exchange. If it is not possible to demonstrate whether it was involved in the exchange then it will be unfeasible to recognize whether its participation was correct. It is straightforward to demonstrate that when verifiable, the TTP cannot be transparent. Therefore, both properties are **incompatible**.

Both properties introduce advantages (see above) in their own context (electronic commerce applications and situations in which the TTP's behaviour cannot be completely trusted). Unfortunately, coexistence of these scenarios is common and as demonstrated, both properties are incompatible to provide by a certified electronic mail application. After analyzing both properties, we assume that the advantage introduced by a transparent TTP is less clear. When a TTP is involved exceptionally with a determined user, there should not be a reason to damage its reputation. Only when a TTP acts continuously due to a determined user's participation should the system (scheme) damage its reputation. From this discussion, we could assert that an external reputation system as defined and used nowadays in different applications and protocols

(e.g. eBay) seems a more natural solution for keeping honest users reputation than a transparent TTP. On the other hand, achieving verifiability in the TTP's protocol involvement seems more important, since this property allows dealing with unfair situations.

5. COMMUNICATION CHANNELS AND TIMELINESS

Communication channels and timeliness are two aspects that are strongly correlated with. This means that when designing a certified electronic mail protocol, both characteristics are to be explicitly defined since they will influence the context in which the application can be used. Generally, three different types of communication channels can be defined:

Definition 16 (Operational channel). A communication channel is said to be operational if, and only if, messages arrive correctly to the recipient in a finite and known amount of time.

It is not realistic to assume this kind of channels in heterogeneous networks as Internet. Under some protocols it is feasible to assume the message will eventually arrive to its recipient. Unfortunately, it is more difficult to achieve the temporal restriction in existing networks. Therefore, the following communication channels need fewer restrictions:

Definition 17 (Unreliable channel). A communication channel is said to be unreliable if, and only if, messages can be permanently lost.

This is the channel with fewest restrictions, but at the same time it greatly determines the properties the CEM (and in general the fair exchange) application will be able to provide. Obviously, this type of communication channel is not unrealistic: a communication channel can be broken (voluntarily or not); a user's computer can be infected by trojans or viruses which avoid the computer network communication; the computer can be subject to a denial of service attack, etc. Tacitly, messages can be lost in the network. Nevertheless, we claim that there exist protocols to recover from these situations such that messages are not permanently lost (can be recovered for instance with replay and acknowledge messages, computer viruses disinfection, network links redundancy, etc.). This leads us to the following definition:

Definition 18 (Resilient channel). A communication channel is said to be resilient if, and only if, messages arrive correctly to the recipient in a finite but unknown amount of time.

That is, attacks and network errors are possible as in unreliable channels but the situation will be eventually corrected and messages retransmitted, thus arriving to the recipient in a finite amount of time. Since we do not know whether retransmissions will occur due to network errors (or attacks), the final amount of time to reach its destination is unknown. Although the unreliable channel imposes fewer restrictions, it will greatly influence the complexity needed to design a valid and fair certified electronic mail protocol.

Note that in the previous definitions there is always a temporal parameter. It is thus appropriate to correlate the communication channel model with other property:

Definition 19 (Timeliness). A CEM protocol fulfils timeliness if, and only if, honest entities can unilaterally choose to terminate the protocol in a finite amount of time without losing fairness [OZL09].

In order to understand this property's importance, we should analyze protocols that do not satisfy timeliness. Some approaches impose deadlines in order to finish or start some steps of the protocol (e.g. a deadline for the protocol execution or a deadline before or after which the TTP can be contacted, etc.). This type of protocols introduces a minor and a serious disadvantage. The former one assumes entities can synchronize their clocks. This clock network synchronization is not trivial, but feasible by means of existing synchronization protocols. A major disadvantage appears because, a priori, the use of deadlines is not compatible with unreliable or resilient channels since it could be a danger for the fairness property (which, as explicitly stated, is a compulsory property). Beforehand, deadlines use is only acceptable under operational communication models (at least between users and the TTP).

Nevertheless, here we must distinguish between different types of deadline time. For instance, Micali presented a CEM scheme [Micali03] in which the deadline indicates the point before which the recipients must contact the TTP if needed, because afterwards, it will not help, and thus, fairness will be damaged (see [OZL09] for further reference).

Other solutions use the deadline as point of end for the protocol (i.e. moment in time from which only the TTP can provide evidences). It can be argued that, at the end, both kind of solutions need a point of time in which the TTP cannot help any more and if recipients do not access in advance, fairness will be also damaged. Nevertheless, there is a very important difference in both approaches which will be perhaps better appreciated in the next figure.

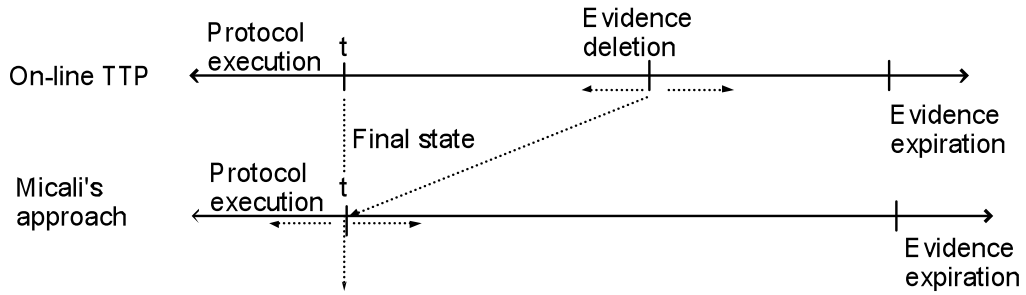


Fig 1. Timeline.

In these solutions, when evidence deletion occurs at the TTP, the latter will not be able to help entities any more. This event can coincide with evidence expiration or not, depending on TTP's storage resources. Nevertheless, time (deadline t) at which the state is final (synchronous timeliness) and evidence deletion are different. In this way, entities can know beforehand when the protocol has a definitive state and then, if needed, seek TTP's help. This type of trade-off can be considered because the use of deadlines is common in current applications, even if the considered channels are not operational. For instance, most protocols use local clocks in order to acknowledge or retransmit messages.

On the other hand, those solutions (as that from Micali) which unifies deadline and evidence deletion (there is not deletion event, but the meaning is the same; i.e., TTP's help is not available) are not possible without considering an operational channel.

Additionally to the importance of timeliness property in CEM applications, we can introduce different variants of this property attending to the existence of deadlines in the protocol itself [OZL09]:

Definition 20 (Synchronous timeliness). A CEM protocol is said to respect synchronous timeliness if all honest entities are able to terminate the protocol in a finite and known amount of time without losing fairness.

In this case, deadlines are used and the TTP clock is assumed as the reference time (i.e., users' clocks need to be synchronized or negligibly unsynchronized with TTP's clock). Though more difficult for users, the TTP does not need to maintain evidences for long-time periods. Hence it can be used with a stateless TTP (see next section for further reference).

Definition 21 (Asynchronous timeliness). A CEM protocol is said to respect asynchronous timeliness if all honest entities are able to terminate the protocol at any time without losing fairness.

In this case, there are no deadlines for participants in the protocol, but a serious practical implication makes it hard to achieve: for this property to be fulfilled, an infinite state (or at least until evidence expiry date if it is the case) has to be maintained by the TTP until parties fetch evidence (stateful TTP). Otherwise, if responsible of distributing evidence, it needs to retry till recipients of evidence acknowledge reception. An alternative solution can be to extend the channel reliability between TTP and users (in this direction).

However, those solutions that do not impose any temporal restriction (**definition 21**) do not force any channel property and thus, they are most appropriate. Nonetheless, there will be a trade-off between the design complexity and the efficiency needed by the CEM application in order to settle which timeliness property will be finally implemented (always preserving fairness).

In most solutions, it is assumed that channels among users are unreliable and that the channel with the TTP is resilient. When possible, only one-way resilient channel is assumed; i.e., the user will eventually contact the TTP but this does not imply that the latter will eventually contact the user. Although we defined operational channels as well, they make very strong assumptions on the network services and are rather unrealistic in heterogeneous networks.

6. STATE STORAGE

A traditional classification criteria for application servers is whether they store state information. Since TTP is implemented as an application server it can be classified by this criteria. Hence, TTPs can be classified with respect to how long (temporal criteria) do they need, if applicable, to store state information.

Definition 22 (Strong stateless TTP). A TTP is said to be a strong stateless TTP if, and only if, it can resolve all requests from all users without storing previous request information.

Obviously, this is, from a resource and storage capacity point of view, a desirable property to be fulfilled by the TTP. Nevertheless, forcing the TTP (i.e. the application implementing its

functionality) to resolve every request without previous received information could lead to complex and inefficient CEM protocol and application designs.

Definition 23 (Weak stateless TTP). A TTP is said to be a weak stateless TTP if, and only if, in order to resolve all requests from all users it can access previous exchange information, but the latter could be deleted after a finite and known amount of time.

If the participating entities agree on a deadline for finishing the certified email exchange, the TTP will not need to store the corresponding information after this deadline. It is also possible that the TTP could delete the state information if all participating entities (sender and receiver) have already contacted it (for instance, because due to protocol design no more TTP interaction is feasible). In any case, the TTP needs to guarantee that no entity loses fairness due to state information deletion.

Definition 24 (Strong stateful TTP). A TTP is said to be a strong stateful TTP if, and only if, in order to resolve all requests from all users it needs to access previous exchange information which will be stored forever.

The typical scenario for the need of this property is that in which the TTP's state information (stored evidences or signed requests) might be needed in case of a dispute resolution process between sender and receiver. This is because that unless otherwise specified by the protocol or application policy, there will be no deadline for a dispute resolution process between participants. Even though it is not clear what will be the value of a certified email exchanged after a long period of time, this cannot be quantified.

Definition 25 (Weak stateful TTP). A TTP is said to be a weak stateful TTP if, and only if, in order to resolve all requests from all users it can access previous exchange information, but the latter could be deleted after a finite and unknown amount of time.

This is the usual type of TTP implemented in optimistic approaches fulfilling the asynchronous timeliness property **(definition 21)**. As users can (repeatedly) send requests to the TTP at any time, responses need to be coherent with those previously provided.

7. EFFICIENCY

There is no doubt that efficiency is a parameter which needs to be considered when evaluating CEM protocols. Together with security properties this is a determinant factor when comparing and evaluating different approaches. Nevertheless, it is not trivial to determine what an efficient solution is and even more difficult to consider which is the optimal one. As in other services in the telecommunications field (and the certified electronic mail is a telecommunications' service) we should firstly define a metric in order to compare different solutions.

Definition 26 (Efficiency). A solution A is more efficient than a solution B if, and only if, under the same hardware and network resources, A is executed faster than solution B (in average).

In general, it can be considered that the cost of operations needed to execute the protocol (message's generation and especially cryptographic computations) is much lower than that of transmitting the messages in the different steps of the protocol. Thus, a priori, we can set that a protocol that needs N different steps (in each step there is a transmission of message(s)) is more efficient than a protocol that needs $N+1$ steps. The first problem turns up because the number of steps executed by a protocol is not always the same (e.g. optimistic protocols which run different sub-protocols in case of exceptions). In this case, we can evaluate the best and worst case as is commonly done in the Algorithm's Analysis (in short, a protocol is implemented as an algorithm).

Bearing this in mind, we introduce three important claims for the efficiency property:

Claim 1. A fair CEM protocol for two entities A and B with two-steps does not exist.
Proof: Clearly, a two-steps protocol is impossible to develop a fair exchange. If A sends her element to B, B possesses the element from A and now he can decide not to send his item to A.

Claim 2. A fair CEM protocol for two entities A and B with three-steps does not exist if the TTP acts in-line or on-line.
Proof: If the TTP acts in every execution of the protocol, it will act, in its simplest way, as a relay. This means for an exchange to succeed, a minimum of 4 steps are needed. A sends its item to the TTP which in turn relays it to B and vice versa. Indeed, this is the simplest fair-exchange protocol with an in-line TTP: A and B send their items to the TTP (2 steps) which then forwards B's item to A and A's item to B (2 steps).

Claim 3. A fair CEM protocol for two entities A and B with three-steps exists if the TTP acts off-line and moreover, it is optimal.

Proof: Once we have demonstrated Claim 2, we just need to show an existing fair optimistic CEM protocol with only three steps (in the best case). This has been published in [FPH00].

Finally, it is important to highlight that efficiency is just a key property for CEM protocols and it cannot be the only property considered to evaluate the overall convenience of an application. Indeed, if a protocol is very efficient but due to its design it does not fulfil a desired security property it will be rather a bad option. Therefore we stress that in order to completely evaluate a solution all desired and all security properties need to be assessed.

8. DISPUTE RESOLUTION AND EVIDENCE TRANSFERABILITY

For a CEM protocol to be practical it needs to provide digital evidence to the users such that they can demonstrate whether the exchange occurred (and under what conditions). This is very important because after the exchange, sender and addressee can dispute whether the exchange existed or, for instance, what was the content of this particular exchange. Typically there are two possible situations:

- **repudiation of receipt**: The sender alleges having sent a certified message while the recipient denies receiving it.
- **repudiation of origin**: The addressee alleges receiving a message while the sender denies having sent it.

Almost all existing approaches allow resolving both kinds of disputes suitably. Those solutions with sender's anonymity[3] make an exception because, as explained in section 9, they are incompatible with non-repudiation of origin. The existing dispute resolution processes differ in the participation of the different actors to the process. Generally the dispute resolution is driven by an external entity. In some solutions, originators and recipients can turn to the arbiter separately and the latter can resolve the dispute only taking into account the evidence provided by the claiming party. Nevertheless, other solutions are designed such that the arbiter can only deterministically resolve the dispute interrogating (requesting evidence) the other participants in the protocol (in some cases even the TTP).

Beforehand, it seems natural the fact that the external arbiter could need to request proving evidence to some or all of the participating entities in the protocol. If one party claims that an

[3] When the anonymity provided as in the case of [KM01] is only temporal during the protocol execution in order to fulfil no author-based selective receipt, NRO can be provided.

exchange occurred and the other party denies it (or vice versa), it seems reasonable for the arbiter to examine both parties' evidence before settling the final report. In specific situations also the TTP records could be examined, as it participated in the protocol. Actually we made clear that also this entity can misbehave and thus can be charged, and it needs to be able to defend itself. Nonetheless, some actors could refuse to participate in the dispute resolution process or it could be difficult to get them involved because of several other reasons. Thus, in principle, autonomous resolution processes seem to be a better option. Additionally, other reasons favour this type of solutions:

Definition 27 (Evidence transferability). A CEM protocol generates transferable evidences if, and only if, after the protocol run, sender and recipient can separately demonstrate to any third party whether the message was finally sent/received without the need to request other entities' input.

This property is important not only because the user's autonomy to dispute whether a message was sent/received, but also because we can find use cases in which this property is relevant. Let's assume we receive a certified notification stating that we are the winners of a lottery game. It will be for us very interesting to be able to use this notification in order to demonstrate to a financial institution that we have sufficient balance to order a money transfer, without the need for the institution to request additional evidence to sender and/or TTP about the authenticity and integrity of the evidence received. Furthermore, this property is very useful and generally studied by authors in other applications as, for instance, contract signing protocols. If this property is considered as relevant for a CEM application, those protocols which allow entities to receive contradictory evidences are to be discarded.

We would like to highlight that this property is not directly related with fairness, that is, a protocol can generate non-transferable evidence and be completely fair (obviously it will need the intervention of additional participating entities in case of disputes). Some solutions [FPH00] have been erroneously criticized as unfair because they did not generate transferable evidences.

9. CONFIDENTIALITY AND ANONYMITY

Message content confidentiality is not an intrinsic requirement in certified mailing (and thus in CEM applications). Each user needs to decide which information is sensible and use the correct mechanisms to keep it confidential. In all existing protocols the message content is initially protected (e.g. encrypted) in order to avoid the recipient accessing the content before providing

evidence of receipt. After this moment, keeping the message confidentiality is only an optional property.

When this property needs to be fulfilled, it is desirable to keep the message secrecy also when a TTP is involved in the exchange:

Definition 28 (Confidentiality). A CEM protocol fulfils the confidentiality property if, and only if, for the correct and fair protocol execution no third entities need access to the message content.

When this property is referred to the TTP, some authors call it as **neutral** TTP whereby other authors designate it as an **operational** TTP. Our definition is not limited to the TTP, but to any other entity (participating or not in the protocol or in a possible dispute resolution process) different from sender and receiver. Evidently, if any of this entity acts as a notary, this property would be inappropriate.

In this section we must stress that this property is optional and it depends on the application of the certified electronic mail service. Therefore, a CEM solution cannot impose this property into the design of the protocol itself.

The originator's privacy, generally, is also optional and it depends on the application itself. However it is important to stress that originator's anonymity and NRO cannot be provided at the same time (incompatible properties). We can demonstrate it as follows:

Proof: A non-repudiation of origin and anonymity properties are incompatible in any general communication protocol. The demo can be easily inferred from **definition 2**. If the non-repudiation of origin property requires the generation of a digital evidence in order to determine unambiguously whether the sender is the message's author such that it cannot deny having sent a message, regardless of the mechanism used to achieved this property, this will prevent the originator from being anonymous.

10. FULFILMENT OF PROPERTIES.

In the introduction of the paper, properties are classified as mandatory, optional or desirable. However, there are some incompatibilities among properties that avoid the fulfilment of all the desirable properties in the same protocol. We finish the description of the properties summarizing these incompatibilities in table 2 and showing in table 3 the sets of properties fulfilled by some CEM protocols.

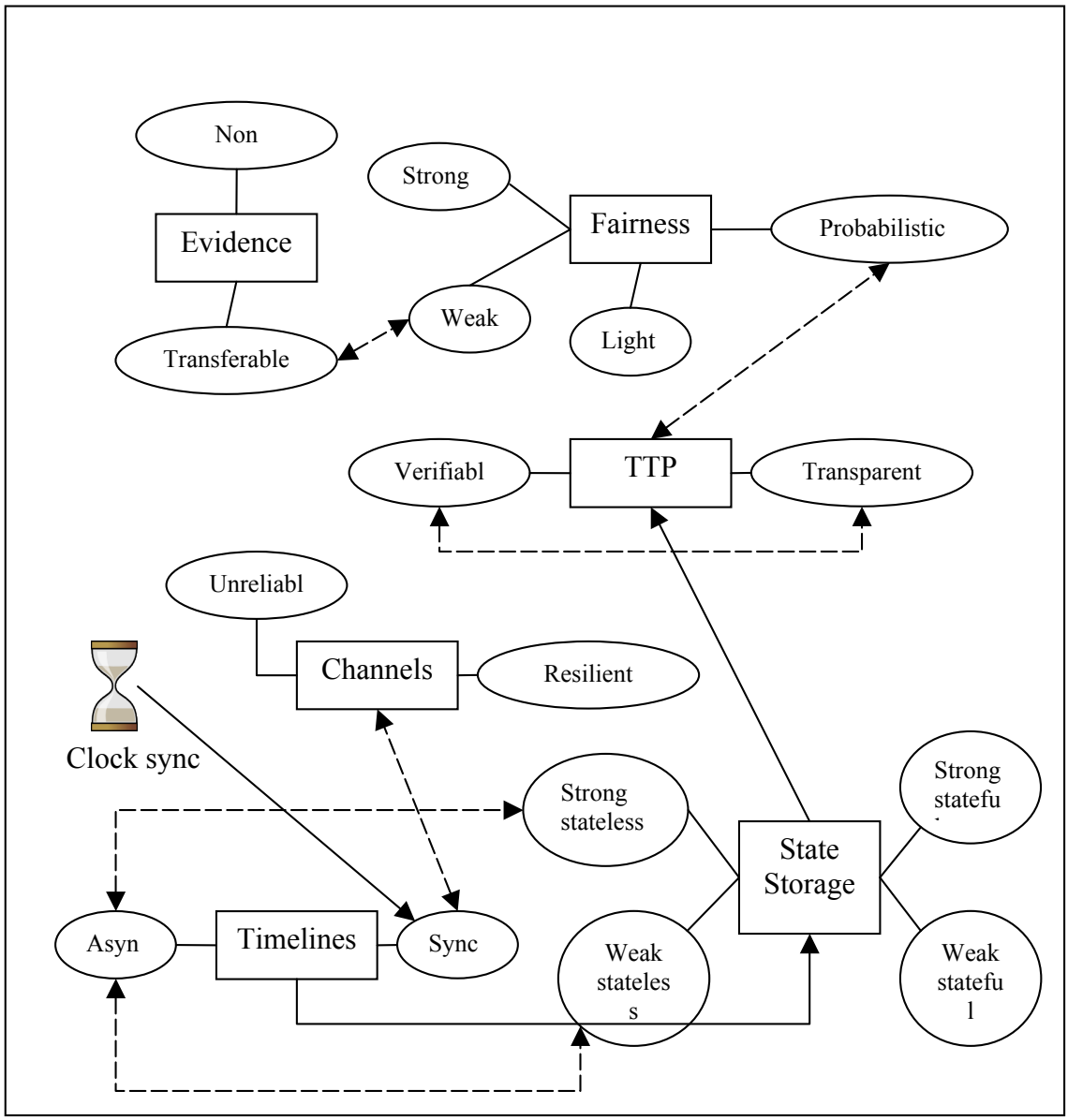
Incompatible Properties	Description of the Incompatibility
1. **Transparent TTP vs. Verifiable TTP**	The proofs generated by a transparent TTP cannot be distinguished from the proofs generated by the parties without the intervention of the TTP. For this reason, it can not be proved that the TTP has been involved in the execution of the protocol. If the involvement can not be proved, it is evident that the correctness of its behavior could neither be proved, so the TTP can not be verifiable.
2. **Timeliness vs. Strong Stateless**	If only one party can contact the TTP: The TTP will decide the final state of the exchange. After that, the TTP will contact the other party to send him the result. It is not necessary to store any information. But the second party must wait a possible message from the TTP and he can not know the final state of the exchange at any moment, so the exchange is not timeliness. If both parties can contact the TTP: a synchronous solution can allow the synchronization of the requests, but an asynchronous solution requires the storage of information because this information can change the final state of the exchange. If the information would not be stored the exchange could be unfair.
3. **Timeliness vs. Weak Stateless**	If only one party can contact the TTP, the exchange is not timely terminated. If both parties can contact the TTP, and the TTP temporarily stores some information, then the parties will have a deadline to contact the TTP (before the deletion of the information), so the protocol would not be fulfill asynchronous timeliness.
4. **Transferable proofs vs. Weak fairness**	Those solutions that allow the generation of contradictory evidences or that require the request of the evidences to both parts in order to know the final state of the exchange cannot transfer the proofs to external arbiters. As a consequence, only the protocols that achieve strong fairness provide transferable proofs.
5. **Synchrony vs. Unreliable and Resilient Channels**	Synchronous solutions require the finalization of the exchange before a deadline. But resilient and unreliable channels do not guarantee the delivery of the messages before a deadline. Synchronous solutions combined with resilient channels can derive in unfair situations.

Table 2. Incompatible Properties

Incompatibility 5 needs further attention. As theoretically stated in the table above, it is not possible to reach synchrony when using unreliable and/or resilient channels. Nevertheless, some protocols [OZL04a, ZOL05, ZG96, Zhou01] defines synchronous timeliness with this type of channels. This is the case of [OZL04a1].

These protocols make the practical assumption that the messages will arrive to the TTP before the deadline established using a resilient channel, thus allowing the participating entities to bring the protocol to completion at an unknown time after the deadline. Note that even if this practical assumption does not hold, fairness is not damaged (only timeliness). This is because the deadline defines the time *after* which sender and addressee can contact the TTP to retrieve information (see **definition 19** for further details). Nowadays, some protocols, even if implemented over unreliable channels (i.e. Internet) use deadlines for flow control, protocol termination, protocol refresh, etc. For synchronization with the TTP clock, different protocols exist over unreliable channels (e.g. Network Time Protocol, Precision Time Protocol for local arena networks, etc.).

In the next figure, we can see a graphical representation of the properties presented in this paper:



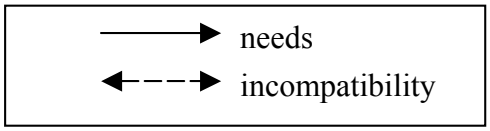


Fig 2. Dependencies Graph.

	[OZL04a1]	[MR99]	[PEC,PICT]	[FPH00]	[KM01]	[OZL04a2][4]
Fairness						
Strong			X		X	
Weak				X		
Light	X					X
Probabilistic		X				
TTP						
In-line			X			
On-line	X					
Off-line				X	X	X
Transparent						(X
Verifiable	X			X		X)
Communication Channels and Timeliness						
Resilient	X			X	X	X
Unreliable	X		X		X	X
Operational		X				
Synchronous	X					
Asynchronous		X		X	X	X
State Storage						
Strong stateless						
Weak stateless	X		X			
Strong stateful						
Weak stateful				X	X	X

[4] The protocol can be configured to fulfil either transparency or verifiability.

Efficiency						
3-steps best case				**X**		**X**
Anonymity						
No author-based selective receipt					**X**	
Evidence Transferability						
Transferable	**X**	**X**	**X**		**X**	
Non-transferable				**X**		**X**

Table 3. Fulfilment of Properties.

11. CONCLUSIONS

In this paper we claim that there exists no consensus with respect to the definition of the certified electronic mail service, and that this lack extends even to the properties that the service needs to fulfil in order to be considered a useful digital service. Unfortunately, the research community has not reached a general common view. For instance, although an agreement exists in the opinion that certified electronic mail should involve a fair exchange of items, there is no concurrence with respect to the items to be exchanged.

Precisely, in this paper we present a number of definitions with the main goal of providing a common understanding on the issue, focusing our attention on revisiting the properties of CEM. This helps to elucidate and refine a number of significant issues that researchers still find quite fuzzy when investigating in the existing literature. Our contribution is of great benefit for the design and development of suitable and homogeneous CEM applications. In this sense, we elaborate on some definitions related to fairness because even this property has different approximations for different authors. Then, we comment on the participation of trusted third parties in CEM applications, as well as on the TTP transparency and verifiability, also properties of interest in such applications.

Communication channels and timeliness are two strongly correlated aspects, what means that when designing a CEM protocol both characteristics need to be explicitly defined given that they influence the context in which the application can be used. Hence, this is also part of our focus in the paper. Additionally, state storage and efficiency are, as well, parameters that need to be considered when evaluating CEM protocols. Both are determinant factors when comparing

and evaluating different approaches and, for this reason, we have included them as an important part of our research.

In order for a CEM protocol to be practical, digital evidences are compulsory so that users can demonstrate whether the exchange occurred. The importance is justified because, after the exchange, sender and addressee can dispute about the existence and content of the particular transaction. Therefore, some important definitions on this are introduced in our work. Finally, confidentiality and anonymity of message content, even if not being intrinsic requirements of certified mailing, are also considered such that each user decides which information is sensible for that kind of specific protection.

BIBLIOGRAPHY

[AG02] Abadi, M. & Glew, N., Certified email with a light on-line trusted third party: design and implementation Proceedings of the eleventh international conference on World Wide Web, ACM Press, 2002, 387-395 .

[Ateniese01] Ateniese, G.; de Medeiros, B. & Goodrich, M.T. TRICERT: A Distributed Certified E-Mail Scheme Network and Distributed System Security Symposium Conference Proceedings, 2001, 47-56.

[Aso98] Asokan, N.: Fairness in electronic commerce. Ph.D. thesis, University of Waterloo, Computer Science (1998)

[DGLW96] Deng, R.; Gong, L.; Lazar, A. A. & Wang, W. Practical Protocols For Certified Electronic Mail Journal of Network and Systems Management, 1996, 4, 279-297.

[ETSI] ETSI TR 102 605 v1.1.1 (2007-2007) Technical report. Electronic Signatures and Infrastructures (ESI); Registered E-mail.

[FPH02] Ferrer-Gomila, J.L.; Payeras-Capellà, M. & Huguet-Rotger, L. A Realistic Protocol for Multi-Party Certified Electronic Mail. Information Security ISC 2002, LNCS 2433, 210-219.

[FPH00] Ferrer-Gomila, J.L.; Payeras-Capellà, M. & Huguet-Rotger, L. An Efficient Protocol for Certified Electronic Mail. ISW '00: Proceedings of the Third International Workshop on Information Security, LNCS 1975, 237-248.

[KMZ02] Kremer, S.; Markowitch, O. & Zhou, J. An intensive survey of fair non-repudiation protocols, Computer Communications, 2002, 25, 1606-1621.

[KM01] Kremer, S. & Markowitch, O., Selective Receipt in Certified E-Mail, Advances in Cryptology: Proceedings of Indocrypt, 2001, LNCS 2247, pp. 136-148, Springer-Verlag.

[LNJ00] Liu, P.; Ning, P. & Jajodia, S. Avoiding Loss of Fairness Owing to Process Crashes in Fair Data Exchange Protocols. Proceedings of 2000 International Conference on Dependable Systems and Networks, 2000, 631-640 .

[Micali97] S. Micali. Certified e-mail with invisible post offices. Available from author: an invited presentation at the RSA'97 conference, 1997.

[Micali03] Micali, S. Simple and fast optimistic protocols for fair electronic exchange Proceedings of the twenty-second annual symposium on Principles of distributed computing, ACM Press, 2003, 12-19.

[MR99] Markowitch, O. & Roggeman, Y. Probabilistic Non-Repudiation without Trusted Third Party Second Workshop on Security in Communication Network~99, 1999.

[Nicolas05] Nicolás González-Deleito, No Author-Based Selective Receipt in Certified Email with Tight Trust Requirements, 4th International Workshop on Applied PKI (IWAP 2005). Pages 78–91. September 2005. IOS Press.

[Oppliger04] Oppliger, R., Certified mail: the next challenge for secure messaging Commun. ACM, ACM Press, 2004, 47, 75-79.

[OZL04] Onieva, J. A.; Zhou, J. & Lopez, J. Enhancing Certified Email Service for Timeliness and Multicast. Fourth International Network Conference, 2004, 327-335.

[OZL04a] Onieva, J. A.; Zhou, J. & Lopez, J. Non-repudiation Protocols for Multiple Entities. Computer Communications, 2004, 27, 1608-1616.

[OZL09] Onieva, J. A.; Zhou, J. & Lopez, Secure Multi-Party Non-repudiation Protocols and Applications, Series: Advances in Information Security , Springer, Vol. 50, 2009, ISBN: 978-0-387-75629-5.

[PEC] Posta Elettronica Certificata.
http://www.cnipa.gov.it/site/it-IT/Attività/Posta_Elettronica_Certificata__(PEC)/

[PFR05] Puigserver, M.; Ferrer Gomila, J. & Rotger, L. Certified e-Mail Protocol with Verifiable Third Party e-Technology, e-Commerce and e-Service, 2005. EEE '05. Proceedings. The 2005 IEEE International Conference on, 2005, 548-551.

[PIT] Poste Italiane. http://www.poste.it/online/postemail/

[WBZ04] Wang, G.; Bao, F. & Zhou, J. On the Security of a Certified E-Mail Scheme Lecture Notes in Computer Science, Progress in Cryptology - INDOCRYPT 2004: 5th International Conference on Cryptology in India, 2004, 3348, 48-60.

[Zhou01] Zhou, J. Non-repudiation in electronic commerce, Artech House, 2001.

[ZG96] Zhou J. and Gollmann D. Certified electronic mail. In Proc. ESORICS '96, Lecture Notes in Computer Science, 1996, 1146, 160-171.

[ZOL05] Zhou, J.; Onieva, J. A. & Lopez, J. Optimised Multi-Party Certified Email Protocols. Information Management & Computer Security Journal, 2005, 13, 350-366.