Idea: A Unifying Theory for Evaluation Systems

Giampaolo Bella[1] and Rosario Giustolisi[2]

[1] Dipartimento di Matematica e Informatica, Università di Catania, Italy
giamp@dmi.unict.it,
[2] IT University of Copenhagen, Denmark
rosg@itu.dk

Abstract. Secure systems for voting, exams, auctions and conference paper management are theorised to address the same problem, that of secure evaluations. In support of such a unifying theory comes a model for *Secure Evaluation Systems* (SES), which offers innovative common grounds to understand all four groups. For example, all rest on *submissions*, respectively votes, test answers, bids and papers, which are to be *evaluated* and ultimately ranked. A taxonomy for all groups is advanced to provide a comparative understanding of the various systems. The taxonomy is built according to the type of submissions and the type of evaluation.

The uniformity of the security requirements across all groups offers additional validation, and this is an innovative finding in the direction, currently unexplored, of a common system design. Still, the requirements may variously shape up. For example, while voter privacy is normally required forever, anonymity of the submissions is required until after the marking/evaluation phase for the test answers of an exam, for the (sealed) bids of an auction, and for the papers submitted to a conference.

1 Introduction

There are at least four groups of secure systems that are widely used at present. These are respectively for *voting*, *exams*, *auctions* and *conference paper management*. Each group has been extensively studied so far. To advance an example system per group, we mention Helios for voting [1], Remark! for exams [2], the protocol presented by Curtis et al. for auctions [3] and Confichair for conference paper management [4].

This idea paper unfolds our theory that all groups can be unified at an abstract level. The theory is supported by three main pillars. One is a formal model for *Secure Evaluation Systems* (SES), whose main elements are the submitters, the authorities, the submissions and an evaluation function (§2). The model is a tuple that can be instantiated over each group or a specific system, thus offering a benchmark for a contrastive assessment of the various systems.

Another pillar in support of our unifying theory is a taxonomy for the groups of systems based upon the type of submissions and the type of evaluation (§3). For example, the taxonomy supports the claim that exam systems and conference systems are very similar, although only exams may seek submissions of type ordered choice, namely a ranked list.

The third pillar is a requirement elicitation process across the four groups of systems (§4). It is found that all systems have in common various flavours of authentication, non-repudiation, fairness and privacy. In particular, receipt-freeness and coercion-resistance, traditionally spelled out for voting, are interpreted for the first time for exams, indicating the impossibility for an examinee to prove the ownership of her test until after the marking, even with the collaboration of a coercing examining authority. By contrast, after the marking terminates, the system should allow the examinee to publicly leverage the mark for her test.

2 A Model for Secure Evaluation Systems

Secure Multi-Party Computation (SMPC) is a widely studied area of cryptography aimed at the distributed, privacy preserving computation of a function [5]. It means that all participating players will provide inputs that are needed to compute the function, whose output may be made public; however, the computation must not reveal anything about the inputs, hence preserve the privacy of the players. This model can be reviewed to emphasise the details of secure evaluation systems. It is useful to further detail our four groups of secure systems.

Voting system is a method for making a decision or expressing an opinion, usually following discussions, debates or election campaigns. The submissions consist of a set of preferences (votes) over some options (candidate, decisions, etc.). The evaluation consists of a tallyng algorithm that outputs a ranking of candidates (or a winning candidate).

Exam system is a method for evaluating candidates according to their knowledge or skill. The submissions consist of a set of tests over some options (open-ended questions, multiple-questions, etc.). The evaluation consists of a marking algorithm that outputs a ranking of tests (or a winning test).

Auction system is a method for buying and selling goods or services by offering them up for bidding, then taking the bids, and finally selling to the winning bidder. The submissions consist of a set of offers (bids) over some options (goods, prices, etc.). The evaluation consists of an algorithm that outputs a ranking of bids (or a winning bid).

Conference system is a method for managing the papers to be presented at a conference and often published in a book of proceedings. The submissions consist of a set of papers, which are often anonymised. The evaluation consist of an algorithm that outputs a ranking of papers (or a winning paper).

This description underlines clear similarities among all groups, such as that they all aim at producing a ranking. However, the evaluation used for the raking is inherently different. While there is no notion of "correctness" of a vote in democracy, there clearly is such a notion for test answers. Also, while all bids are potentially correct once they are in the right format, correctness of a research paper also is meaningful.

An informal definition can be given to identify the subject matter.

Definition 1 (SES — Informal). *A* Secure Evaluation System *is a SMPC system that computes a function termed* evaluation function *securely. Its players can be partitioned as* submitters*, who contribute* submissions*, and* authorities*, who contribute* administration*.*

A formal model can then be built to capture a SES abstractly. The model rests on a set S of *submitters* a set s of *submissions* and a set A of *authorities*. The players treat the submissions by means of a set T of *tasks*, such as sending the submissions or entering data in a computer. The specific list of tasks is normally understood as the protocol definition underlying the system. The tasks may express important features of a SES, for example as an electronic protocol if the tasks occur over computing devices, or as a face-to-face protocol if the tasks take place traditionally, *de visu*. Both submitters and authorities may misbehave to obtain an advantage maliciously. This admits a threat model, namely a set of malicious tasks T_t.

The evaluation function f, which may be jointly computed by the players, should satisfy a set R_f of functional requirements. The privacy preservation prerequisite can be generalised as a set R_s of security and privacy.

A SES can thus be formalised as a tuple.

Definition 2 (SES — Formal). *A* Secure Evaluation System*, at the formal level, is a tuple* $SES = \langle \mathsf{S}, \mathsf{A}, \mathsf{s}, \mathsf{T}, \mathsf{T}_\mathsf{t}, \mathsf{f}, \mathsf{R}_\mathsf{f}, \mathsf{R}_\mathsf{s} \rangle$ *such that:*

- S *is a set of* submitters*;*
- A *is a set of* authorities*;*
- s *is a set of* submissions*;*
- T *is a set of* tasks*, which the players carry out;*
- T_t *is a* threat model*;*
- f *is an* evaluation function*, which the players may jointly compute;*
- R_f *is a set of* functional requirements*;*
- R_s *is a set of* security requirements*.*

The model can be easily instantiated over a target secure system of our four chosen groups. We instantiate it over the groups themselves, building a table that expresses an inclusion relation, Table 1. Therefore, the table is incomplete because it only provides a limited set of examples, but offers a compact, unifying workbench. This highlights a minor ambiguity in the terminology, that a candidate in voting is someone who can be voted for, while a candidate in exam is someone who is examined.

It must be emphasised that all groups of systems are aimed at computing a ranking. This underlines the competitive nature of the problems that all systems address. Also the R_s line is limited, providing just one obvious security requirement per group, but a more comprehensive analysis will follow (§4).

3 A Taxonomy for Secure Evaluation Systems

We build a taxonomy based upon the types of submissions and the type of evaluation. Submissions can be of three types.

$\subseteq$	**Voting**	**Exam**	**Auction**	**Conference**
S	voters	candidates (examinees)	bidders	authors
A	talliers, officials	invigilators, examiners	auctioneer	program chair
s	votes	test answers	bids	research papers
T	vote casting	answering questions	bidding	paper writing
$\mathsf{T_t}$	voting twice	over-marking	bid alteration	de-anonymisation
f	candidate ranking	test ranking	bid ranking	paper ranking
$\mathsf{R_f}$	efficiency	efficiency	efficiency	efficiency
$\mathsf{R_s}$	voter privacy	anonymous marking	bid sealing	anonymous reviewer

Table 1. An incomplete demonstration of the SES formal model

3.1 Types of submissions

Single-choice submission allows the submitter to select one of the possible options. In voting, this submission type reflects Single-Mark Ballot type where each voter chooses one candidate. In exams, it reflects both open-ended tests and those multiple-choice tests that only demand one answer. In auctions, it reflects Dutch and Sealed first-price auction types where each bidder may only put in one bid. In conferences, this type of submissions is the standard one.

Check-All-That-Applies (CATA) submission allows the submitter to select more than one of the possible options, precisely all those that the submitter deems appropriate. In voting, this reflects Approvals ballot type where each voter can select any number of candidates of her choice. In exams, it reflects tests with more than one correct answer. In auctions, it reflects English auction types where bidders can submit multiple bids to get the standing bid. In conferences, it may be interpreted as the submission of more than one paper by the same author list.

Ordered-choice submission allows the submitter to order the options according to a stated criterion. In voting, this submission type reflects Rank and Score ballot types where each voter produces a hierarchy of the candidates. In exams, it reflects scale format tests, where submissions are based on a rating scale. In auctions, it reflects Combinatorial auction type where each bidder can place bids on combinations of discrete items. In conferences, this type of submissions does not seem to be used.

3.2 Types of evaluation

The tasks for evaluating the submissions managed through a SES may, in turn, be carried out in three alternative ways, depending on who performs them (while meeting the requirements in $\mathsf{R_s}$):

Authority evaluation prescribes the submissions to be evaluated by a set of dedicated authorities.

- **Peer evaluation** sees the evaluation of the submissions being performed by a set of peers of the submitter's. Also in this case, anonymity may contribute to the submitter's trust in the peers for the sake of evaluation; for example, the submissions might be anonymised. Additionally, the evaluation may extend, as prescribed by $\mathsf{R_f}$ of the specific protocol, to the answers of a subset or all of the submitters, as we shall see below.
- **Self evaluation** limits the evaluation to be carried out by the individual submitter, namely each individual can perform the evaluation that meets the requirements stated in $\mathsf{R_f}$. Soundness and fairness of such an evaluation are not obvious so will have to derive from the specific tasks of the system.

3.3 Taxonomy

With all the details provided above, a taxonomy can be built for secure evaluation systems. The taxonomy is in Table 2: a cell mentions a group of secure systems when we are aware that there exists at least one system in the group that exhibits the specific combination of submission and evaluation types that the cell pinpoints.

		Evaluation		
		Authority	**Peer**	**Self**
Submission	**Single-choice**	voting auction exam conf.	exam conf.	voting auction
	CATA	voting auction exam conf.	exam conf.	voting
	Ordered-choice	voting auction exam	exam	

Table 2. A taxonomy for Secure Evaluation Systems

Authority evaluation. Most SES's feature an authority that takes care of the evaluation process. All democracies elect holders of offices by voting systems that have tallying authorities. This applies to single-choice, CATA, and ordered-choice types of submissions. In electronic voting, some systems have been proposed to distribute the trust among different authorities. For example, one such system is Helios [1]. Similarly, in most classic auctions, the auctioneer is the authority who declares the winning bid. E-bay is a popular example of electronic auction with a CATA type of submission. The auction system by Curtis et al. [3] fits any submission type. In entrance examinations, authorities normally produce the list of admitted candidate. Some effort to distribute the trust on such authorities has been discussed in a recent proposal of a secure exam system [2]. Notably, the latter fits any submission type. In Easychair, the program chair acts as authority and decides the list of accepted papers. Easychair accepts submissions of more than one paper by the same author list, hence supports both single-choice and CATA submission types.

Peer evaluation. Peer evaluation is peculiar to conferences and exams. For example, in MOOCs homeworks are peer-reviewed. To our knowledge, neither voting nor auction systems have been proposed so far with this feature.

Self evaluation. Kiayias and Yung [6] introduced the notion of self-tallying voting for single-choice submission type, in which the result can be tallied and verified by anybody. Hao et al. [7] proposed a different system that supports an approval ballot type, hence a CATA submission. No exam systems today support ordered-choice with self evaluation. Recent works on smart contract technology, such as AuctionHouse [8] seems to lead to auctions with self-declaration of winning bids enforced by the use of blockchains. To our knowledge, there is no work on exam with self-evaluation, although the use of smart contracts may favour the construction of such a kind of systems.

4 Requirement Elicitation for Secure Evaluation Systems

We now wonder whether it is also possible to find similarities among the SES multi-objective security goals. More specifically, given any security goal of a group of SES systems, can we find a similar interpretation in each of the other groups? History of secure systems tell us that it is unfeasible to list and freeze all the security goals of a system because people's needs may change over time, hence the system's requirements tend to change as well. However, while we may not reach a definitive answer to our question, we may find a temporary answer by considering the main requirements that are popular nowadays. In our analysis, we focus on classic authentication, non-repudiation, fairness, and privacy goals.

Authentication. Data origin Authentication naturally maps to the authentication of the submissions of a SES system. Data origin authentication is a common goal with the same interpretation in voting, exams, auctions, and conferences. It is normally expected that any evaluation algorithm considers only inputs submitted by eligible parties: only ballots cast by eligible voters should be recorded in a voting system; only test answers originated with eligible candidates should be marked in an exam; only bids put by registered bidders should be considered in an auction; only papers by registered authors should be considered as valid submissions to a conference.

In the same way, data origin authentication is also expected for authenticating the outcome of the evaluation in each of the systems. It means that the rankings in all four groups of systems are generated by the corresponding set of official authorities. Note that data origin authentication does not imply the *correctness* of the evaluation, which means that the outcome derives by correct execution of the evaluation function. Data origin authentication guarantees that such a function is fed with all and only eligible submissions. However, correctness of the evaluation is a desired goal for each of our systems, and has a similar interpretation across each of them.

Non-repudiation. An interpretation of non-repudiation [9] is the impossibility for submitters to deny their participation. In voting, it means that a voter cannot deny to have participated in an election. The same clearly applies to exams with candidates, to auctions with bidders and to conferences with papers. However, auctions support an additional interpretation in which non-repudiation may signify the impossibility for a bidder to claim that she did not submit the winning bid. A similar interpretation is hard to find in voting, in which the very opposite is actually desirable, namely that a voter cannot prove the way she voted (receipt freeness). We observe that in exams a test should eventually be linked to the corresponding author in order to assign a mark to each examinee, hence non-repudiation applies to exams.

Another instantiation of non-repudiation regards the reception of submissions. This interpretation applies to voting, exams, auctions and conferences, so that no authority can successfully deny having received valid submissions.

Fairness. As regards submissions, fairness means that choices are submitted independently from other submissions. In voting, it means that no voter can be influenced by votes already cast. In most auction types, submitted offers should not influence subsequent offers. However, this interpretation of fairness obviously does not apply for English auctions, in which bidders submit new offers to displace the standing bid. From a bidding strategy point of view, if we consider the submission of a bidder as the final bid she wishes to offer for an auction, we can see our fairness interpretation in English auctions as well. Fairness is of utmost importance in exams and means that candidates should answer their test based on their knowledge and skills. An additional fairness goal exists for exam and can be named *marking fairness*: it prescribes that tests should be marked independently from the identity of their authors. Note that marks are not the outcome of exam's function evaluation but they rather are inputs to the function to calculate the rank of the tests. Marks can be associated to weights in voting and auctions, in which votes or bids have different weights. Such interpretation, however, requires weights to depend on the identity of the submitters. Thus, an interpretation similar to marking fairness is hard to find in voting and auctions. By contrast, fairness in conferences abides by the same interpretation as for exams.

Privacy. Privacy goals have seen many interpretations. If we look at the privacy of the submission, the interpretation in voting is that the system does not reveal how a voter voted. The same applies to the pairs examinee/test, bidder/bid and author/paper. Note that this definition is strongly related to the definition of fairness discussed above, and the same considerations made about English auctions apply here. Voting systems normally require vote privacy to hold even after the evaluation. The winning bid is normally revealed in auctions, still the identity of the bidder may not be disclosed. The same applies for exams in which the right to publicly disclose the link of a test with its author is left to the examinee. By contrast, this link is routinely disclosed in conferences, where the author is normally required to attend and present the accepted paper. Further differences among those systems find a place in specific definitions of privacy. Strong privacy definitions in voting state that a voter cannot prove the way she voted (*receipt-freeness*) even if the voter collaborates with the coercer (*coercion-resistance*) [10]. Similar strong privacy definitions are less meaningful in auctions since winning bids are normally announced publicly. Also information revealed through other channels, such as who is the (new) owner of the auctioned good or service, would disclose if a bidder sent a winning or a losing offer.

In exams, receipt-freeness and coercion-resistance are meaningful through the marking phase and can be seen as stronger definitions of marking fairness. In particular, receipt-freeness and coercion-resistance are two detailed instances of *anonymous marking*, in which tests are marked while ignoring their authors. They signify that an examinee should not be able to prove the ownership of her test until after the marking (receipt-freeness) even if the examinee collaborates with the coercer, e.g. the examiner (coercion-resistance). However, the possibility of a covert channel between examinee and examiner should be ruled out. Privacy over conferences can be interpreted much the same way as with exams.

Although we found many similar security goal interpretations among SES systems, there may still be differences in other clusters, such as verifiability and accountability [11]. The requirement elicitation needs to be expanded also over such clusters to fully substantiate a putative claim that all systems state the same security requirements.

5 Conclusions

Secure systems for voting, exams, auctions and conferences have never been analysed comparatively before. Our unifying theory claims that this is possible, and our supporting model confirms their similarities. Our taxonomy favours a comparative understanding of the various systems. The traditional security requirements of authentication, non-repudiation, fairness and privacy apply to all four groups. The next step is to focus on either one of the three introduced pillars, i.e., formal model, taxonomy, or requirement elicitation, and to study it thoroughly. For example, it would be interesting to study those group of systems, such as surveys, that normally do not produce a ranking.

A readily-exploitable value of this work is a deep understanding of the security requirements of each system; this is made possible precisely by their argumentation across the various groups. An additional value is that it may inspire a combination of the research efforts that are currently spent in each individual group towards solving more effectively than before what seems to be a same problem, that of secure evaluation.

Acknowledgement

This work is supported in part by DemTech grant 10-092309 from the Danish Council for Strategic Research, Programme Commission on Strategic Growth Technologies.

References

1. Adida, B.: Helios: Web-based open-audit voting. In: Proceedings of the 17th Conference on Security Symposium. USENIX Symposium (2008)
2. Giustolisi, R., Lenzini, G., Ryan, P.Y.: Remark!: A secure protocol for remote exams. In: Security Protocols XXII. Volume 8809 of Lecture Notes in Computer Science. Springer International Publishing (2014) 38–48

3. Curtis, B., Pieprzyk, J., Seruga, J.: An efficient eauction protocol. In: Proceedings of the The Second International Conference on Availability, Reliability and Security, ARES. ARES, IEEE Computer Society (2007) 417–421
4. Arapinis, M., Bursuc, S., Ryan, M.: Privacy supporting cloud computing: Confichair, a case study. In: Proceedings of the First International Conference on Principles of Security and Trust. POST'12, Berlin, Heidelberg, Springer-Verlag (2012) 89–108
5. Yao, A.C.: Protocols for secure computations. In: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science. SFCS, IEEE (1982) 160–164
6. Kiayias, A., Yung, M. In: Self-tallying Elections and Perfect Ballot Secrecy. Springer Berlin Heidelberg, Berlin, Heidelberg (2002) 141–158
7. Hao, F., Kreeger, M.N., Randell, B., Clarke, D., Shahandashti, S.F., Lee, P.H.J.: Every vote counts: Ensuring integrity in large-scale electronic voting. USENIX Journal of Election Technology and Systems (JETS) (2014) 1–25
8. Petkanics, D., Tang, E.: Auctionhouse. `http://auctionhouse.dappbench.com/` (2016) Accessed: 2017-01-16.
9. Kremer, S., Markowitch, O., Zhou, J.: An intensive survey of fair non-repudiation protocols. Comput. Commun. (2002) 1606–1621
10. Delaune, S., Kremer, S., Ryan, M.: Coercion-resistance and receipt-freeness in electronic voting. In: 19th IEEE Computer Security Foundations Workshop (CSFW'06). (2006) 12 pp.–42
11. Küsters, R., Truderung, T., Vogt, A.: Accountability: Definition and relationship to verifiability. In: Proceedings of the 17th ACM Conference on Computer and Communications Security. CCS '10, New York, NY, USA, ACM (2010) 526–535