Co-Utility for Digital Content Protection and Digital Forgetting

Josep Domingo-Ferrer
Universitat Rovira i Virgili,
UNESCO Chair in Data Privacy,
Department of Computer Engineering and Mathematics,
Av. Països Catalans 26,
E-43007 Tarragona, Catalonia
E-mail: josep.domingo@urv.cat

David Megías
Universitat Oberta de Catalunya,
Internet Interdisciplinary Institute (IN3),
Estudis d'Informàtica, Multimèdia i Telecomunicació,
Av. Carl Friedrich Gauss, 5
E-08860 Castelldefels, Catalonia
E-mail: dmegias@uoc.edu

***Abstract*—Tracing the flow of content on the Internet is a difficult problem, due to the distributed nature of transactions and the lack of a common legal framework. In order to address two related problems, digital content protection and digital forgetting, we introduce the co-utility concept. A protocol is said to be co-utile if helping the other peers that participate in the protocol to increase their utilities is the way to increase one's own utility. We then present peer-to-peer protocols that leverage co-utility to provide rationally sustainable solutions for the two above-mentioned problems.**

Keywords: Co-utility, Multicast fingerprinting, Anonymous fingerprinting, Digital forgetting, Game theory.

I. Introduction

In our information society, everyone can store a lot of information in a perfect way and at a very low cost. In fact, vast quantities of data (big data) about everything and everyone are continuously being gathered and digitally stored. Stored data supplement our failing human memory, so that we can remember a lot and other people can remember a lot about us. This may seem an advantage when *we* are the ones who remember, but it is clearly a shortcoming when others remembers things about us that we would prefer to be forgotten. Perfect remembering may not even be good when we are the ones who remember: *e.g.* remembering wrongs too well for our entire lifetime may be a hindrance to our happiness. For the above reasons, the right of an individual to enforce oblivion for pieces of information about her is increasingly being regarded as part of her fundamental right to privacy. For example, the forthcoming General Data Protection Regulation [1], which is about to be adopted by the European Union, enshrines the right to be forgotten. However, the only technical solution currently used to enforce this right is to exclude the items to be forgotten from the results returned by search engines: clearly, this does not mean the information is effectively forgotten (*e.g.* any small change in an item to be forgotten will cause the search engines to return the item again).

A related problem is the necessity of legal distribution of multimedia contents. However, copyright infringement in the redistribution of contents is a practice that has become the most important threat to the electronic commerce industry. Anonymous fingerprinting is a technological solution to prevent illegal content redistribution. Fingerprinting consists of inserting an embedded and imperceptible mark (the fingerprint) in the distributed content in order to identify the content buyer. In case of illegal redistribution, the fingerprint allows identification of the illegal redistributor through a tracing system. This facilitates taking legal actions against copyright infringement. In this solution, law-abiding buyers should stay anonymous and their identity should only be revealed in case of illegal redistribution.

Regarding the content distribution itself, if the content is to be distributed to a group of N receivers, one option is for the content sender to engage in N unicast transmissions, one for each intended receiver, and another option is a single multicast transmission to the entire group. Certainly, the multicast option has the advantage of being faster and more bandwidth-efficient from the sender's point of view. However, the unicast approach has the strong point of allowing the sender to fingerprint the content sent to each receiver by embedding a different serial number in each sent copy, with the aim of detecting and tracing unlawful redistribution of the content. Fingerprinting can additionally be anonymous, so that the sender does not learn the identity of receivers unless they become unlawful redistributors (see [11], [2] for early anonymous fingerprinting proposals). Note that the multicast approach does not allow fingerprinting, as all receivers obtain exactly the same content. Hence, the unicast approach, in spite of its inefficiency, seems more suitable when the sender is a merchant selling content and the receivers are buyers.

Peer-to-peer (P2P) distribution of content appears as a third option blending some of the advantages of the unicast and multicast solutions. P2P distribution of all types of files and contents has become extremely popular with the increased bandwidth of home Internet access in the last few years. In addition, P2P file sharing applications are not restricted to this use, and some companies are also exploiting the P2P distribution paradigm as a way of saving server bandwidth and speeding up the downloads of their products (such as multimedia contents and software updates). Indeed, when using a P2P network for content distribution, the merchant only needs to establish direct connections with one or a few seed buyers, say $M \ll N$ buyers, and send them copies. The content is further spread over the P2P network by those seed buyers. The challenge is how to ensure that the P2P-spread content is still traceable in case of redistribution.

Contribution and plan of this paper

In Section II, we introduce the co-utility notion, which refers to protocols in which mutual help is the best rational option to take, even for purely selfish agents. In Section III, we review several digital content protection and digital oblivion P2P protocols which are rationally sustainable because they are co-utile. To justify their co-utility, we present novel game-theoretic analyses of those protocols. Conclusions are drawn in Section IV.

II. Co-utility

We need P2P protocols for digital content protection and digital oblivion to be such that participating peers are willing to follow them, that is, they ought to be self-enforcing. Yet, we go one step beyond and we will seek self-enforcing P2P protocols that also promote mutually beneficial collaboration between peers, that is, a collaboration that improves the utilities of the involved agents with respect to a non-collaborative setting. To capture the idea of mutually beneficial self-enforcing collaboration, we propose the notion of *co-utility* [6], [5]. Specifically, a protocol is *co-utile* if helping the other peers that participate in the protocol to increase their utilities is the way to increase one's own utility. In the rest of this section, we define co-utility is a more precise way using game theory, which is a natural framework to deal with rational agents that are utility maximizers.

The kind of scenarios to which we apply co-utility can be represented as perfect-information games, that are games in which each agent making a decision knows the payoffs of all agents under the various possible actions (or sequences of actions), plus any previously made decisions [7]. This is coherent with the usual sequential nature of protocols, in which the current state is known. Specifically, we represent these games in the so-called extensive form, which is a tree where: (i) nodes are the points where decisions are made, (ii) each node is labeled with the name of the agent making the decision, (iii) outgoing edges in a node represent the available choices (actions) at that node, and (iv) each leaf node is labeled with the tuple of payoffs that agents obtain when the node is reached.

By using this extensive form, we can view a protocol execution (i.e., the actions needed for the completion of a task) as a path that traverses the tree representing the game. If the protocol allows random choices between actions, we identify the protocol with a subtree rather than a path.

We focus on self-enforcing protocols, which are those from which agents have no rational incentive to deviate. That is, no agent can increase her utility by deviating from the protocol, provided that the other agents stick to it. In game-theoretic terms, this means that, at each successive node of the protocol path, sticking to the next action prescribed by the protocol (taking the next edge in the path) is an *equilibrium* of the remaining subgame of the game (the subtree rooted at the current node), that is, a *subgame perfect equilibrium* of the game.

We say that a self-enforcing protocol is *co-utile* if it results in mutually beneficial collaboration between the participating agents. More specifically, a protocol P is co-utile if and only if *the three* following conditions hold:

1) P is self-enforcing;
2) The utility derived by each agent participating in P is strictly greater than the utility the agent would derive from not participating;
3) There is no alternative protocol P' giving greater utilities to all agents and a strictly greater utility to at least one agent.

The first condition ensures that, if participants engage in the protocol, they will not deviate. The second condition is needed to ensure that engaging in the protocol is attractive for everyone. The third condition can be rephrased in game-theoretic terms by saying that the protocol is a Pareto-optimal solution of the underlying game.

In some scenarios, we can even reach a stronger case of mutual collaboration in which each agent achieves her *maximum* utility in the game. In this case, the protocol does not only lead to a Pareto-optimal payoff assignment to players, but it gives maximum payoff to all players. We call this *strict co-utility*.

In general, given any game, there is no guarantee that the selfish behavior of the players will lead to co-utility, let alone strict co-utility. However, for a perfect-information game where there is one and only one leaf node that strictly maximizes the utility of all agents, the selfish behavior of the agents will cause them to follow a protocol that is strictly co-utile.

See [5] for an example of the concepts we have introduced in this section. In the following sections, we discuss co-utile P2P protocols applied to digital content protection and digital oblivion.

III. Co-utile digital content protection and digital oblivion schemes

This section provides an overview of different digital content protection and digital oblivion schemes that are co-utile. In all these schemes, the rational co-operation of the participants leads to maximum utility for all of them, so in fact strict co-utility is attained.

Domingo-Ferrer and Megías [4] proposed a P2P protocol for distributed multicast of fingerprinted content in which cryptographic primitives and a robust watermarking technique are used to provide the buyers with different marked copies of the content. This work uses the reward and punishment concepts of game theory to guarantee that the buyers co-operate within the P2P distribution system to embed the fingerprint and distribute the content. A similar idea is used in [3], where digital oblivion is also introduced by means of expiration dates. The game-theoretic models used in both cases are detailed below.

In [8], [9], all buyers can obtain different fingerprinted copies of the content, but the embedding algorithm needs to be executed only for a few (M) seed buyers. The other buyers obtain their copies by recombining the fragments of the contents obtained from different parent buyers. There are two types of co-operation required among peer buyers for such protocols to effectively protect against unlawful content redistribution:

- *Unique fingerprints*. This co-operation among buyers requires that no buyer transfers her complete copy of the content to another buyer, in order for the fingerprint in the copy held by each buyer to be unique. We argue that such co-operation is rationally sustainable and in fact is strictly co-utile. If a buyer Alice transfers her entire copy to another buyer Bob, Alice may be accused of any illegal redistribution performed by Bob with his copy. Symmetrically, if Bob obtains her entire copy from Alice, Bob may be accused of any illegal redistribution performed by Alice with her copy. Hence, the optimum action for both Alice and Bob is to co-operate to keep their fingerprints unique, as intended in the protocol.
- *Tracing*. Buyers are also supposed to co-operate in tracing any illegal redistributors. The tracing algorithm starts from the seed buyers and searches the distribution graph following a path of "probable" ancestors of the traced fingerprint. These probable ancestors of the traced buyer are required to return the decryption of a hash of the fingerprints of their "children". A correlation function between the fingerprints is also used to determine the most probable ancestor of the next step, until the search finally identifies the traced buyer. If a buyer refuses to compute the correlation, she can be accused using the hash of the fingerprint revealed with the help of one of its parents. Hence, co-operation is strictly co-utile: given the choice between returning or not returning the correlation, returning it gives maximum utility to the buyer.

Finally, in [10] significant improvements to the recombination approach of the previously cited works were proposed. A first improvement is to replace some trusted parties (proxies) of the system by non-trusted counterparts, thus making the approach much more practical. The second relevant improvement is related to the tracing algorithm. The involvement of other buyers in the tracing algorithm, though effective, is somewhat inconvenient for several reasons. First, a buyer may lose her copy of the content due to perfectly legitimate reasons (e.g. hardware damage), which makes it impossible to follow the tracing path beyond a specific node of the distribution graph (another path would need to be explored, if there is any other left). Also, contacting several innocent buyers during the search of the buyer who has infringed copyright can be cumbersome. Hence, the proposed system replaces the graph search by a standard database search of encrypted fingerprints. Hence, co-utile co-operation of buyers is only needed to ensure unique fingerprints, but not for tracing.

The following sections provide a detailed review of these systems, summarize their relevant features and specify game-theoretic models of the players' utilities that justify the co-utile nature of the proposed protocols.

A. Distributed multicast of fingerprinted content based on a rational peer-to-peer community

The system presented in [4] works in the following way:

- The first buyer of the content (P^1) and the merchant (P^0) engage in an anonymous fingerprinting scheme protocol such that P^1 obtains her fingerprinted copy D_{01} and the merchant obtains a transaction record $t_{0,1}$, which can be used later on to retrieve the fingerprint and of P^1 in case of illegal redistribution of D_{01}.
- The next buyers P^i for $i > 1$ engage in the same anonymous fingerprinting scheme in such a way that P^{i+1} obtains the fingerprinted copy $D_{01\dots(i+1)}$ of the content, and P^i obtains a transaction record $t_{i,i+1}$ that is forwarded to the merchant (P^0).

When an illegally redistributed copy is found, the system searches the transaction records from $t_{N-1,N}$ backwards. When a fingerprint is successfully extracted, the identity of the illegal redistributor is revealed.

Game theory is applied in this scheme to enforce "rational" buyers to follow the protocol. The idea is to guarantee that the best strategy for each player (buyer or merchant) is to follow the protocol; in other works, following the protocol yields the maximum possible payoff. More precisely, consider each possible choice of player P^i regarding the distribution protocol:

s^0: The only possible strategy for P^0 is to deliver the content to P^1 by engaging in anonymous fingerprinting with her.

s_0^i: Follow the protocol and engage in anonymous fingerprinting with P^{i+1} and return $t_{i,i+1}$ to P^0.

s_1^i: Deviate from the protocol by engaging in anonymous fingerprinting with P^{i+1} but not returning $t_{i,i+1}$ to P^0.

s_2^i: Deviate from the protocol by not engaging in anonymous fingerprinting with P^{i+1} but returning a fake $t_{i,i+1}$ to P^0.

s_3^i: Deviate from the protocol by not engaging in anonymous fingerprinting with P^{i+1} and not sending any $t_{i,i+1}$ to P^0.

Under s_2^i and s_i^3, P^i sends her copy to P^{i+1} without any fingerprinting. The following costs, payoffs, rewards and punishments are introduced:

d_i: Payoff for P^i as a consequence of obtaining $D_{01\dots i}$ and preserving her privacy.

$-v_i$: Cost (negative payoff) incurred from engaging in anonymous fingerprinting with P^{i+1}.

$-w_i$: Cost (negative payoff) incurred from returning $t_{i,i+1}$ to P^0.

$r_{i,i+1}$: Reward (positive payoff) obtained by P^i for engaging in anonymous fingerprinting with P^{i+1}.

$-p_i$: Punishment (negative payoff) incurred by P^i in case of not sending a valid $t_{i,i+1}$ to P^0 and risk being accused of illegal redistribution.

With this model, the utility function for each strategy can be computed as follows:

$$\begin{aligned} u_i(s_0^i) &= d_i + r_{i,i+1} - v_i - w_i, \\ u_i(s_1^i) &= d_i + r_{i,i+1} - v_i - p_i, \\ u_i(s_2^i) &= d_i - v_i - p_i, \\ u_i(s_3^i) &= d_i - p_i. \end{aligned}$$

Then, the assumptions $r_{i,i+1} \geq v_i$ and $-p_i \leq -w_i$ yield the following ordering for the utilities of the different strategies:

$$u_i(s_2^i) \leq u_i(s_3^i) \leq u_i(s_1^i) \leq u_i(s_0^i).$$

Hence, the maximum utility yields a *dominant strategy* consisting in all players following the protocol correctly, or

$$(s^0, s_0^1, \ldots, s_0^{N-1}, s_3^N).$$

Although the system is explained as if the distribution was just a line from P^0 to P^N, it can be easily transformed into a tree-distribution protocol in which each buyer engages into several further distributions with other buyers.

B. Rational enforcement of digital oblivion

The approach taken in [3] is similar to that of [4], but an expiration date is also embedded into the original content D_0 to guarantee digital oblivion. If T_0 is the expiration date of the content, P^0 embeds it into D_0 using a blind and robust watermarking algorithm such that T_0 can be recovered by anyone who receives the content, but it cannot be removed from D_0 without a substantial damage in the content quality. After this first step, the remaining part of the distribution protocol is similar to that of [4].

In this scenario, the objective of the system is to detect public posting of the content beyond the expiration date T_0. Hence, all copies of the content used in the public domain shall be removed by the expiration date. Otherwise, actions can be taken against the buyer of the content who would be identified by means of her embedded fingerprint.

The game-theoretic model is similar to that of the multicast distribution system, but there is a slight difference in the strategies of buyers. s_0^i, s_1^i, s_2^i and s_3^i are the same as in Section III-A and there is the following additional strategy:

s_4^i: Deviate from the protocol by not forwarding the content and not sending any $t_{i,i+1}$ to P^0.

In this case, the payoffs, costs, rewards and punishments are as follows: d_i, $-v_i$, $-w_i$, $r_{i,i+1}$ and $-p_i$ defined in the same way as for the multicast distribution scheme of Section III-A, but we have also a negative payoff (communication cost) $-f_i$ that P^i incurs from anonymously forwarding the unfingerprinted content (or, more precisely, the content without the corresponding additional fingerprint) to P^{i+1}. In this case, we have $f_i \leq v_i$, since forwarding the content is less costly than engaging into anonymous fingerprinting with P^{i+1}.

The utilities for the different strategies can be computed as follows:

$$\begin{aligned} u_i(s_0^i) &= d_i + r_{i,i+1} - v_i - w_i, \\ u_i(s_1^i) &= d_i + r_{i,i+1} - v_i - p_i, \\ u_i(s_2^i) &= d_i + r_{i,i+i} - f_i - w_i - p_i, \\ u_i(s_3^i) &= d_i + r_{i,i+i} - f_i - p_i, \\ u_i(s_4^i) &= d_i. \end{aligned}$$

The following assumptions are then made: $r_{i,i+1} \geq f_i$, $v_i + w_i < p_i$ and $r_{i,i+1} \geq v_i + w_i$ which yields $u_i(s_1^i) \leq u_i(s_0^i)$, $u_i(s_2^i) \leq u_i(s_0^i)$, $u_i(s_3^i) \leq u_i(s_0^i)$ and $u_i(s_4^i) \leq u_i(s_0^i)$. This imples that following the protocol is the best strategy for all buyers in terms of achieving maximum utility.

C. Recombination-based anonymous fingerprinting

The basic features of the fingerprinting scheme proposed in [8], [9] are the following:

- The system consists of a merchant, seed buyers, non-seed buyers, a group of P2P proxies (trusted), a transaction monitor and a tracing authority (trusted).
- The content is divided into several ordered fragments and in each of them a binary sequence, called a "segment", is separately embedded. The concatenation of all segments forms the whole fingerprint.
- The merchant creates a set of M seed copies and distributes them to the M seed buyers in such a way that their corresponding fingerprints have low pair-wise correlation. The embedding process is required only for these M seed buyers.
- Non-seed buyers obtain the fragments of the content from at least two other buyers (see discussion on co-utile unique fingerprints at the beginning of Section III) and her fingerprints are built as a recombination of the segments of her parents, as shown in Fig. 1.
- For the non-seed buyers of the system, different fingerprints are created without any further execution of the embedding scheme.
- An onion routing-like protocol using a P2P proxy (or a chain of proxies) is used within the system to provide anonymous communications between peer buyers. The proxies do not know the real identities of the source and destination buyers.
- The proxies are responsible for forwarding a one-time symmetric session key from the child buyer to the parent buyer in each transaction in order to encrypt the transferred content fragments for data security.
- The transaction monitor keeps a record of all transactions conducted between peer buyers for the transfer of a specific content. These records make it possible to know the specific providers (peers) for each buyer.
- The merchant is the only party who has access to the buyers' database (with real identities).
- In case an illegally redistributed copy is found, the redistributor can be identified using a graph search directed by a binary correlation function between the fingerprint

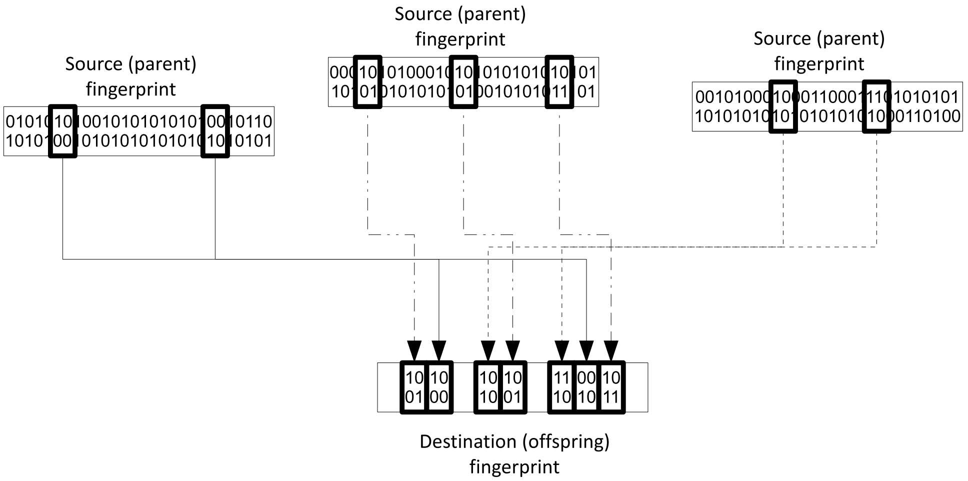

Fig. 1. Automatic recombined fingerprint construction

extracted from the redistributed copy and the fingerprints of the tested buyers.

- In each step of the traitor tracing protocol, the buyer with the maximum correlation is chosen as the most likely ancestor of the illegal redistributor. Occasionally, a higher correlation may be obtained for a non-ancestor of the buyer, which requires exhausting a subgraph and backtracking.
- Cooperation is required from some honest buyers during the graph search process (see discussion on co-utile tracing at the beginning of Section III), though the fraction of tested buyers is shown to asymptotically decrease to zero as the "population" of buyers grows if no backtracking is required during traitor tracing.

As mentioned above, in this scheme co-operation of peer buyers is required in two different protocols: 1) in content distribution to ensure unique fingerprints; and 2) in traitor tracing in case of illegal redistribution. Although no game-theoretic model is given in [8], [9] for the rational involvement of players in these two protocols, in this paper we propose such a model for the distribution protocol.

Let P^i be a source player who is contacted by a proxy P^j to deliver the content to a requesting buyer P^k. The possible strategies for P^i in the distribution protocol are the following:

s_0^i: Follow the protocol and engage in anonymous transfer with P^j to transfer no more than 50% of the fragments encrypted using the symmetric key k provided by P^j.

s_1^i: Deviate from the protocol and engage in anonymous transfer with P^j to transfer more than 50% of the fragments encrypted using the symmetric key k provided by P^j.

s_2^i: Deviate from the protocol and engage in anonymous transfer with P^j without encrypting the fragments.

s_3^i: Deviate from the protocol and do not engage in anonymous transfer with P^j.

As detailed in [8], [9], the fragments are signed from the origin. Consequently, the possibility of sending fake fragments is excluded from the set of strategies for P^i.

Regarding P^j, the proxy is trusted by assumption and can only choose the following strategy:

s_0^j: Follow the protocol and engage in anonymous transfer with P^i first, and then with P^k. Transfer the symmetric key k anonymously from P^i to P^k. Receive the encrypted fragments from P^i and forward them to P^k. Send a report of the transaction to the transaction monitor.

While [8], [9] assume trusted proxies, this is relaxed in Section III-D.

Finally, P^k can choose among the following strategies:

s_0^k: Follow the protocol and engage in anonymous transfer with at least two proxies P^j by requesting no more than 50% of the fragments to each proxy.

s_1^k: Deviate from the protocol and request more than 50% of the fragments of the content to the same proxy P^j.

With these strategies, the following payoffs, costs, rewards and punishments can be defined for these three types of players:

d_k: Payoff for P^k as a consequence of obtaining the content and preserving her privacy.

$-v_i, -v_j, -v_k$:
Costs for P^i, P^j and P^k, respectively, as a consequence of participating in the anonymous transfer protocol.

$-p_i, -p_k$:
Punishments (negative payoffs) incurred by P^i and P^k, respectively, for having a copy of the content with more than 50% of the fragments identical to the fragments of another buyer and risk being unfairly accused of illegal redistribution due that other buyer's misbehavior.

$-w_i$: Cost incurred by P^i for transferring unencrypted segments to P^j.

$-q_i$: Punishment (negative payoff) incurred by P^i for distributing unencrypted fragments such that intermediate routers or other parties can have access to the cleartext of the fragments, which may result in P^i being accused of illegal redistribution.

r_i: Reward (positive payoff) given to P^i for transferring the fragments to P^j.

In this model, we have $w_i < v_i$, since encrypting and sending the fragments (as done in the anonymous transfer) is more costly than sending them in cleartext. The resulting utilities for the three types of players are the following:

$$\begin{aligned}
u_i(s_0^i) &= r_i - v_i, \\
u_i(s_1^i) &= r_i - v_i - p_i, \\
u_i(s_2^i) &= r_i - w_i - q_i, \\
u_i(s_3^i) &= 0, \\
u_j(s_0^j) &= -v_j, \\
u_k(s_0^k) &= d_k - v_k, \\
u_k(s_1^k) &= d_k - v_k - p_k.
\end{aligned}$$

If $v_i < r_i$ and $w_i + q_i > v_i$, the best strategies for all players are the ones that lead to following the protocol correctly: s_0^i, s_j^0, s_k^0. Note, however, that the assumption of honest proxies is required: proxies must engage in the protocol even if they only obtain a negative payoff (cost) from doing so. This trust assumption on proxies is removed in the next section and a more complex game-theoretic model is required in that case.

D. Improved recombination-based anonymous fingerprinting

The recombination-based system presented in [8], [9] has two relevant problems: 1) trusted proxies are required for content distribution and 2) the co-operation of some innocent buyers is required also for traitor tracing. The improved system of [10] overcomes these two drawbacks. First, the distribution protocol is modified using the transaction monitor as a temporary (trusted) key database for the exchange of the symmetric keys between P^i and P^k. Second, the tracing protocol is completely modified in such a way that the graph search is replaced by a standard database search of encrypted fingerprints. The first improvement makes it possible to consider malicious proxies, which means that more strategies are available for the proxies in the game-theoretic model. The second improvement prevents some situations in which illegal redistributors could not be traced due to concealment or abetment by one (or more) of her "ancestors". These reasons may lead to a fully untraceable subgraph of peer buyers.

The new distribution protocol works as follows: 1) P^i chooses the symmetric key k and a pseudorandom binary sequence r to be used as a handle (primary database key) for k. 2) P^i sends (r, k) to the transaction monitor (trusted for key management), who stores it in a database. 3) P^i sends r to P^j, who forwards it to P^k. 4) P^k sends the handle r to the transaction monitor, who replies with the symmetric key k. 5) The transaction monitor blocks the record (r, k) for a given period (timer). When the timer expires, the transaction monitor removes the record from the database. 6) P^i sends the requested fragments, encrypted with k, to P^j. 7) P^j forwards all fragments to P^k, who can decrypt them using k.

Since malicious proxies can participate in this new system, the set of available strategies for them is larger than in the initial recombination-based scheme. The possible malicious behaviors of a proxy P^j are the following:

m_1^j: Deviate from the protocol and try to obtain the decryption k by sending r to the transaction monitor.

m_2^j: Deviate from the protocol and do not report the transaction record to the transaction monitor.

m_3^j: Deviate from the protocol and report a fake transaction record to the transaction monitor. Note that m_2^j and m_3^j are mutually exclusive.

m_4^j: Deviate from the protocol and do not forward the encrypted fragments to P^k.

m_5^j: Deviate from the protocol and do not request the encrypted fragments from P^i. This behavior necessarily implies m_4^j.

All these malicious behaviors $m_1^j, \ldots, m_5^j$ can be detected in the protocol of [10], and will receive their corresponding punishments:

- m_1^j can be detected as detailed in Theorem 1 of [10], entailing a punishment $-p_{j,1}$.
- m_2^j would be detected when another proxy reports the transaction record. Since there are at least two proxies for each transaction, the probability of being detected is very high, entailing a punishment $-p_{j,2}$.
- Similarly, m_3^j would be detected, since no other transaction record would be reported to the transaction monitor. In addition, random checks (included in the protocol of [10]) imply some probability of random detection of fake data in the record. This entails a (likely) punishment $-p_{j,3}$.
- m_4^j and m_5^j would be reported by P^k, since she would not obtain the requested fragments and report the situation to the system. This leads to punishments $-p_{j,4}$ and $-p_{j,5}$, respectively. Typically, $p_{j,4} > p_{j,5}$, since obtaining content fragments and keeping them should be punished more harshly than just ignoring a demand of P^k to obtain fragments.

Apart from these punishments, the following assumptions about costs and rewards can be made:

- P^j wins a reward r_j for following the protocol correctly. This reward should be enough to compensate the proxy for the cost of following the protocol: $r_j > v_j$.
- The cost of following the protocol in case of a combination of behaviors m_1^j, m_2^j and m_3^j is approximately the same of following the protocol correctly: $-v_j$.
- Behavior m_4 prevents some communications to P^j, incurring a cost $-w_{j,4}$ such that $w_{j,4} < v_j$.

- Behavior m_5 prevents even more communications to P^j, incurring a cost $-w_{j,5}$ such that $w_{j,5} < w_{j,4} < v_j$.

TABLE I
POSSIBLE STRATEGIES AND UTILITIES FOR A MALICIOUS PROXY P^j

Strategy: s_x^j	**Behaviors**	**Utility:** $u_j(s_x^j)$
s_0^j	Follow the protocol	$r_j - v_j$
s_1^j	m_1^j	$-v_j - p_{j,1}$
s_2^j	m_2^j	$-v_j - p_{j,2}$
s_3^j	$m_1^j + m_2^j$	$-v_j - p_{j,1} - p_{j,2}$
s_4^j	m_3^j	$-v_j - p_{j,3}$
s_5^j	$m_1^j + m_3^j$	$-v_j - p_{j,1} - p_{j,3}$
s_6^j	m_4^j	$-w_{j,4} - p_{j,4}$
s_7^j	$m_1^j + m_4^j$	$-w_{j,4} - p_{j,1} - p_{j,4}$
s_8^j	$m_2^j + m_4^j$	$-w_{j,4} - p_{j,2} - p_{j,4}$
s_9^j	$m_3^j + m_4^j$	$-w_{j,4} - p_{j,3} - p_{j,4}$
s_{10}^j	$m_1^j + m_2^j + m_4^j$	$-w_{j,4} - p_{j,1} - p_{j,2} - p_{j,4}$
s_{11}^j	$m_1^j + m_3^j + m_4^j$	$-w_{j,4} - p_{j,1} - p_{j,3} - p_{j,4}$
s_{12}^j	$m_2^j + m_4^j + m_5^j$	$-w_{j,5} - p_{j,2} - p_{j,5}$
s_{13}^j	$m_3^j + m_4^j + m_5^j$	$-w_{j,5} - p_{j,3} - p_{j,5}$

With all these considerations, the resulting strategies and utilities for P^j are summarized in Table I. It can be noticed that all strategies, except the one consisting in following the protocol correctly, result in a negative utility to P^j. Hence, the most rational strategy for all players (including also P^i and P^k, whose utilities are the same as those reported in Section III-C) is to follow the protocol.

IV. CONCLUSION

Digital content protection and digital forgetting on the Internet are difficult problems because of the lack of a common legal framework and the distributed nature of transactions. We have reviewed the concept of co-utility and we have shown that co-utile protocols for peer-to-peer digital content protection and digital forgetting are an attractive option. The game-theoretical analyses of the recombination-based anonymous fingerprinting protocols have been introduced here for the first time.

ACKNOWLEDGMENTS AND DISCLAIMER

The following funding sources are gratefully acknowledged: Templeton World Charity Foundation (grant TWCF0095/AB60 “CO-UTILITY”), Spanish Government (projects TIN2011-27076-C03-01/03 “CO-PRIVACY”, TIN2014-57364-C2-1-R/2-R “SmartGlacis” and TIN2015-70054-REDC), European Commission (project H2020 RIA-644024 “CLARUS”), Government of Catalonia (ICREA Acadèmia Prize to the first author and grant 2014 SGR 537). The first author is with the UNESCO Chair in Data Privacy. The views in this paper are the authors’ own and do not necessarily reflect the views of UNESCO or any of the funders.

REFERENCES

[1] General Data Protection Regulation, European Union. https://en.wikipedia.org/wiki/General_Data_Protection_Regulation

[2] J. Domingo-Ferrer, “Anonymous fingerprinting of electronic information with automatic identification of redistributers,” *Electronics Letters*, vol. 34, no. 13, pp. 1303-1304, 1998.

[3] J. Domingo-Ferrer, “Rational enforcement of digital oblivion,” in *Fourth International Workshop on Privacy and Anonymity in the Information Society*. ACM, 2011, pp. 2:1–2:8.

[4] J. Domingo-Ferrer and D. Megías, “Distributed multicast of fingerprinted content based on a rational peer-to-peer community,” *Computer Communications*, vol. 36, no. 5, pp. 542–550, 2013.

[5] J. Domingo-Ferrer, D. Sánchez and J. Soria-Comas, “Co-utility: self-enforcing collaborative protocols with mutual help”, *Progress in Artificial Intelligence*, to appear.

[6] J. Domingo-Ferrer, J. Soria-Comas and O. Ciobotaru, “Co-utility: self-enforcing protocols without coordination mechanisms”, in *Proceedings of the 5th International Conference on Industrial Engineering and Operations Management-IEOM’15*, IEEE, 2015, pp. 1-7.

[7] K. Leyton-Brown and Y. Shoham, *Essentials of Game Theory: A Concise, Multidisciplinary Introduction.* Morgan & Claypool, 2008.

[8] D. Megías and J. Domingo-Ferrer, “DNA-inspired anonymous fingerprinting for efficient peer-to-peer content distribution,” in *Proceedings of IEEE Congress on Evolutionary Computation- CEC’13*, 2013, pp. 2376–2383.

[9] D. Megías and J. Domingo-Ferrer, “Privacy-aware peer-to-peer content distribution using automatically recombined fingerprints,” *Multimedia Systems*, vol. 20, no. 2, pp. 105–125, 2014.

[10] D. Megías, “Improved privacy-preserving P2P multimedia distribution based on recombined fingerprints,” *IEEE Trans. Dependable and Sec. Comput.*, vol. 12, no. 2, pp. 179–189, 2015.

[11] B. Pfitzmann and M. Waidner, “Anonymous fingerprinting”, in *Advances in Cryptology-Eurocrypt’97*, LNCS 1233, Springer, pp. 88-102, 1997.