Tardos fingerprinting codes in the combined digit model

Tardos Fingerprinting Codes in the Combined Digit Model

Boris Škorić[1], Stefan Katzenbeisser[2], Hans Georg Schaathun[3], Mehmet U. Celik[4]

[1]Eindhoven University of Technology, Dept. of Mathematics and Computer Science,
Eindhoven, The Netherlands
[2]Technische Universität Darmstadt, Security Engineering Group,
Darmstadt, Germany
[3]University of Surrey, Department of Computing
Guildford, Surrey, UK
[4]Civolution, Eindhoven, The Netherlands

May 28, 2009

Abstract

We introduce a new attack model for collusion secure codes, and analyze the collusion resistance of two version of the Tardos code in this model, both for binary and non-binary alphabets. The model allows to consider signal processing and averaging attacks via a set of symbol detection error rates. The false positive rate is represented as a single number; the false negative rate is a function of the false positive rate and of the number of symbols mixed by the colluders.

We study two versions of the q-ary Tardos code in which the accusation method has been modified so as to allow for the detection of multiple symbols in the same content segment. The collusion resilience of both variants turns out to be comparable. For realistic attacker strengths the increase in code length is modest, demonstrating that the modified Tardos code is effective in the new model.

1 Introduction

Fingerprinting provides a means for tracing the origin and distribution of digital data. Before distribution of digital content, it is modified by applying an imperceptible fingerprint, which plays the role of a personalized serial number. The fingerprint is usually embedded through a watermarking algorithm. Once an unauthorized copy of the content is found, the identity of at least one guilty user, who participated in the creation of the unauthorized copy, can be identified. The latter can be done using a tracing algorithm, which outputs a list of allegedly guilty users who collaborated to generate the unauthorized copy. This is also known as 'traitor tracing' or 'forensic watermarking'.

Reliable tracing of traitors requires security against attacks that are aiming to remove any personal information from a copy. Collusion attacks, where a group of pirates collude to compare there copies, are a particular threat. As any differences between the copies have to arise from the fingerprint and not the contents, such comparison gives information which can be used to remove the fingerprint.

Coding theory has produced a number of collusion-secure codes (e.g. [2, 10]). As these are only codes, they must be combined with some kind of embedding scheme (or modulation), such as a watermarking system. This can be viewed as a two-layer model [5, 8], where the coding layer encodes user identities to protect against collusion attacks, and the underlying watermarking layer hides the message in the digital contents.

Until now, the development of watermarking schemes and fingerprinting codes has been performed mostly independent of each other; the interface between the fingerprinting code and the watermarking

system has been specified in terms of the *marking assumption* and an attack model which specifies the type of symbol manipulation that the attackers are able to perform. According to the marking assumption, colluders are able to perform modifications only in those content segments where the colluders do not all receive identical information. (These segments are called detectable positions.) The attack model describes the power of the colluders. The commonly used *restricted digit model* only allows colluders to 'mix and match' their copies of the content, i.e. the unauthorized copy is only composed of symbols that the attackers have available. The *unreadable digit model* allows for slightly stronger attacks: Besides mixing the content segments, the attackers can also erase the embedded fingerprint at detectable positions. Under the *arbitrary digit model* the attackers can put arbitrary symbols in detectable positions, while the *general digit model* additionally allows erasures at detectable positions.

However, all these attack models fail to completely capture the properties of the watermarking layer. The mismatch is especially pronounced in the case of spread spectrum watermarks. First, the marking assumption does not always hold, since signal processing attacks are occasionally able to remove a watermark symbol in undetectable positions. Furthermore, signal processing attacks result in symbol errors that seem to match the general digit model at a first glance, but actually the general digit model allows for unrealistically strong attacks. Signal processing can induce the following symbol detection errors:

- If the colluders possess many differently watermarked copies of a segment, they have a good chance of erasing the watermark in that segment.
- Depending on the detector threshold, 'false positive' symbol detections can be induced by adding noise.

These detection errors occur with a certain (low) probability. However, the general digit model allows the attackers a 100% success rate. (As a consequence, efficient code constructions for this model are not known.)

In view of this discrepancy between the potency of actual attacks on the one hand and the general digit model on the other hand, we introduce a new attack model which we call the *combined digit model*. We demonstrate that our model is realistic, and consistent with the use of spread-spectrum watermarking in the watermarking layer. It allows for symbol errors with certain (parametrisable) probabilities, resulting from common attacks in the watermarking layer. We show that an efficient fingerprinting code can be constructed in our attack model, namely a variant of the arbitrary-symbol Tardos code [11]. We analyze the performance of this fingerprinting code.

1.1 Related Work

Several fingerprinting codes were proposed in the past in order to solve the problem of collusion attacks against forensic tracking watermarks. Most notable are the codes proposed by Boneh and Shaw [2] and Tardos [10]. The latter one is a fully randomized binary code that achieves a code length of $m = 100c_0^2\lceil \ln \frac{1}{\varepsilon_1} \rceil$, which is asympotically optimal. (Here c_0 denotes the number of colluders that can be resisted, and ε_1 is the maximum allowed probability of accusing a fixed innocent user.) Furon et al. [3] presented an alternative security proof of the Tardos scheme.

Blayer and Tassa [1], Škorić et al. [12] and Nuida et al. [6] showed how to significantly reduce the constant '100' in the length bound. The paper [11] also provided a construction for non-binary alphabets and showed how to reduce the code length even further by introducing a symbol-symmetric accusation strategy. All these results were derived under the common assumption of the Restricted Digit Model. As noted in [11] the nonbinary Tardos code can also be analyzed in the Unreadable Digit Model, where the colluders may erase fingerprint symbols with 100% success rate in detectable positions. In this case, the required code length is considerably longer, which makes the scheme less practical. It must be noted, however, that the attackers in the Unreadable Digit Model are unrealistically powerful.

Another attack model has been defined by Guth and Pfitzmann in [4], which allows for some attacks against embedded watermarks such as 'mix and match' of the fingerprint symbols that the colluders received. Thus this attacker model is less strong than the one we use; codes for their model, based on the Boneh-Shaw code, can be found in [4, 8].

Xie et al. [13] independently introduced two alternative accusation methods for the Tardos code, as well as an attacker model which allows attackers to perform signal processing attacks on the content (which amounts to occasionally falsely detected fingerprint symbols) as well as mixing of several watermark symbols in one position. Their analysis is purely experimental, and shows that the two accusation methods both perform well, almost identically.

Even though spread-spectrum watermarking is known to provide a certain level of collusion-security in itself, without the need for an additional coding or fingerprinting layer, such solutions scale very poorly in the number of users [8]. Existing results have been limited to simulations up to about 5000 users.

1.2 Contribution and outline

In this paper we introduce an attack model which we call the *Combined Digit Model*. The content owner's WM detector can detect multiple symbols in a content segment. In each segment the attackers may use multiple symbols to create their pirated version (provided that they have observed them, in accordance with the marking condition). Depending on how many symbols they used, there is a probability for each of the symbols that it will not be detected. In addition, the colluders may do a processing attack. We represent this as a small probability that a symbol gets detected which was *not* used in the attack. Simulation results confirm that such an attack model is realistic.

We use two accusation methods which are an extension of the symbol-symmetric accusation method in [11], adapted to the detection of multiple symbols per segment; both methods were concurrently proposed in [13]. We analyze the two accusation methods in a different way than [13]. We focus on a different performance parameter, one that is linked to the minimum code length required to resist c_0 colluders. Furthermore, our study is more analytic. The performance parameter is based on the expectation value of the coalition's accusation sum and on the variance of an innocent user's accusation. These quantities are computed almost completely analytically, with one numerical step at the end. An important benefit of this approach is that it is possible to identify analytically a 'worst case' pirate strategy that forces the content owner to use a long code.

The outline of the paper is as follows. In Section 2 we introduce the Combined Digit Model and provide evidence from simulations that the model adequately captures the essential properties of coalition attacks. In Section 3 we describe the extension of the symmetric q-ary fingerprinting method of [11]. In Section 4 we study the performance of the two accusation methods.

2 The attack model

2.1 Notation

Let Σ be the alphabet of the fingerprinting code, n be the number of users to be accommodated in the system and m the number of symbols in the fingerprint (the number of segments in the content). Furthermore, we denote with ε_1 the probability that one specific innocent user gets falsely accused and with ε_2 the probability that the accusation fails to accuse any guilty user. The distributed codewords can be arranged as an $n \times m$ matrix $\mathbf{X}$, where the j-th row corresponds to the fingerprint given to the j-th user. Let C be a set of colluding users. We denote by c the number of colluders and by $\mathbf{X}_C$ the $c \times m$ matrix of codewords distributed to the colluders. The colluders use a (possibly nondeterministic) strategy ρ to create an unauthorized copy of the content from their personalized copies. The unauthorized copy carries a fingerprint y which depends on both the strategy and the received codewords, i.e. $y = \rho(\mathbf{X}_C)$.

Note that while $X_{ji} \in \Sigma$, the attacked fingerprint y_i cannot be expressed as a symbol in Σ.

2.2 The Combined Digit Model

The proposed Combined Digit Model is based on the following observations:

- Current watermarking schemes offer a considerable level of robustness; however, it is still possible to erase watermarks with a small probability, e.g. due to the addition of noise to the content. Thus,

the code must be able to deal with erasures even in undetectable positions.[1]

- Watermark detectors have a small probability of 'false positives' on the watermarking level, i.e., an attacker may modify content (e.g. by adding noise) such that a watermark is present even though a mark was never actively embedded. Even though the probability of this event is rather small, the occurrence of false positives should be part of the model. It was noted in [9] that even an averaging attack by colluders who all have a '0' can result in detection of a '1'.

- For big coalitions, the colluders have a large number of differently watermarked content segments available. The more symbols they have in a detectable position, the easier it is for them to erase the watermark in the colluded copy. On the other hand, if they use averaging with an insufficient number of different symbols, then they run the risk that multiple symbols get detected.

Traditional fingerprinting codes cannot cope with this extended attack model. We thus introduce the *Combined Digit Model* as follows. During watermark embedding, a fingerprint (a row of $\mathbf{X}$) is embedded in the content. For this purpose, the content is divided into m segments; in each segment one symbol is embedded. The colluders output an object carrying a fingerprint $y = (y_1, \ldots, y_m)$. During the accusation process, a watermark detector is available that returns a score for each symbol $\alpha \in \Sigma$ (e.g. a normalized correlation value). We will write $W_{i\alpha} \in \{0, 1\}$ for the watermark detector response on segment i when the presence of the watermark encoding symbol $\alpha \in \Sigma$ is tested.

The Combined Digit Model is parametrized through a number of different probabilities, representing the power of the colluders:

- We denote with r the probability of the event $W_{i\alpha} = 1$, given that the colluders did not use the symbol α to create y_i. (Either they did not have it or they chose not to use it.) We assume that this probability depends neither on i nor on α.

- Let Ω_i denote the set of symbols present in the i'th column of $\mathbf{X}_C$, and $\omega_i = |\Omega_i|$. Let Ψ_i denote the set of symbols that the colluders use to create y_i, and $\psi_i = |\Psi_i|$. (Necessarily $\Psi_i \subseteq \Omega_i$, $\Psi_i \neq \emptyset$.) We define u_ψ to be the probability of the event $W_{i\alpha} = 1$ for $\alpha \in \Psi_i$. Again, we assume that u_ψ is independent of i and α.

The numbers r and u_ψ depend on the amount of noise that can be introduced by the attackers. The attack model implies that whenever the attackers make use of ψ different symbols in a segment, the detector will trigger on these symbols with probability u_ψ, while the detector will only be triggered with probability r on the other symbols. For $\psi = 1$, the detection probability is close to 1. We take u_ψ as a decreasing function of ψ. The choice of ψ is part of the colluder strategy ρ. (And ψ_i cannot exceed ω_i.)

In the limiting case $r = 0$, $u_1 = 1$, $u_\psi = 0$ (for $\psi \geq 2$) the Combined Digit Model reduces to the Unreadable Digit Model.[2] When multiple symbols are used by the colluders, u_ψ is zero and no symbol is detected by the watermark detector. This is equivalent to erasure. Note that the colluders have free choice (in a detectable position) to put a single symbol or an erasure. Under some circumstances an erasure is actually worse for the coalition than a clearly identifiable single symbol [11].

[1] It was already noted in [10] that the binary Tardos code can easily deal with such noise. The code only has to be made slightly longer.

[2] Reminder: In the Unreadable Digit Model the allowed attacks in a detectable position i are (a) choose any of the symbols in Ω_i, and (b) erasure.

Table 1: Notation used throughout the paper.

Symbol	Meaning
n	number of users
m	number of segments
X_{ji}	watermark symbol of user j in segment i
Σ	the alphabet
q	the numbers of symbols in the alphabet. ($q = \|\Sigma\|$)
$F_{q\kappa}$	prob. density for $\mathbf{p}^{(i)}$ in the 1st step of generating $\mathbf{X}$
κ	shape parameter for $F_{q\kappa}$
$p_\alpha^{(i)}$	Prob$[X_{ji} = \alpha]$ in the 2nd step of generating $\mathbf{X}$
C	the set of colluders
c	the number of colluders. ($c = \|C\|$)
$\mathbf{X}_C$	the part of $\mathbf{X}$ seen by the colluders
$b_\alpha^{(i)}$	number of occurrences of symbol α in i'th column of $\mathbf{X}_C$
y_i	attacked watermark in segment i
ρ	the colluder strategy. $y = \rho(\mathbf{X}_C)$
Ω_i	the set of symbols in column i of $\mathbf{X}_C$
ω_i	number of distinct symbols in column i of $\mathbf{X}_C$; $\omega_i = \|\Omega_i\|$
Ψ_i	the set of symbols used by the colluders to create y_i
ψ_i	number of distinct symbols used to create y_i. ($\psi_i = \|\Psi_i\|$)
$W_{i\alpha}$	detector response for symbol α in segment i
r	Prob$[W_{i\alpha} = 1]$ for $\alpha \notin \Psi_i$
u_ψ	Prob$[W_{i\alpha} = 1]$ for $\alpha \in \Psi_i$, with $\|\Psi_i\| = \psi$
Φ_i	$\{\alpha \in \Sigma : W_{i\alpha} = 1\}$
φ_i	$\|\Phi_i\|$
$\mathcal{A}_j$	accusation of user j; accusation method A
$\mathcal{B}_j$	accusation of user j; accusation method B
$\mathcal{A}_\mathrm{C}$	coalition accusation $\sum_{j\in C} \mathcal{A}_j$
$\mathcal{B}_\mathrm{C}$	coalition accusation $\sum_{j\in C} \mathcal{B}_j$
N_i	$\sum_{\alpha\in\Phi_i} b_\alpha^{(i)}$
P_i	$\sum_{\alpha\in\Phi_i} p_\alpha^{(i)}$
$\langle x \rangle$	1 if the boolean x is TRUE, 0 if x is FALSE

2.3 Empirical justification of the attack model

We briefly present simulation results that corroborate the assumptions we made in formulating the Combined Digit Model. The simulations are based on the model of Zhao *et al.* [14], using Gaussian spread-spectrum watermarking with a non-blind detector. The detector uses the Z statistic as recommended in [14]. Each of the q-ary symbols in the outer (Tardos) code is represented by a random Gaussian signal of length $n = 100$, mean $\mu = 0$, and variance $\sigma^2 = 1/9$. The employed attack was averaging with added uniform noise, identified as the best known attack in [7]. Following [14], distortion was measured by MSE-JND (Mean Squared Error Just-Noticeable-Difference), and the attack was calibrated to give an average normalised MSE-JND of 0.01 per sample. The resulting error rates from simulations with 1000 tests are shown in Figures 1 and 2.

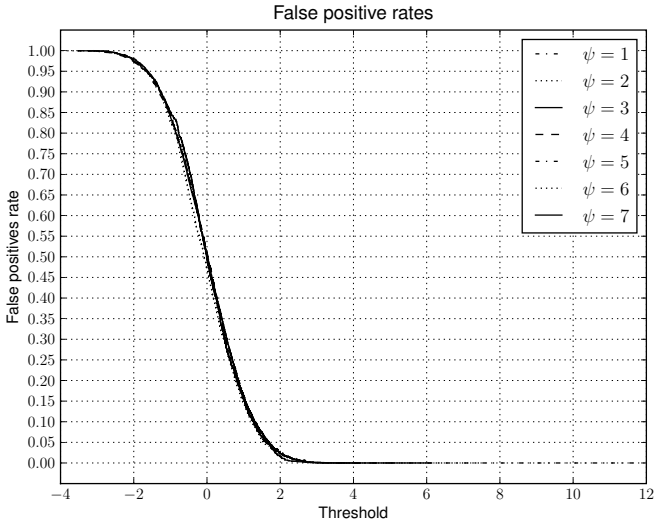

Figure 1: *False positive detection rate r as a function of the detection threshold, plotted for $\psi = 1, \ldots, 7$.*

Fig. 1 shows the false positive rate r as a function of the detection threshold, plotted for several values of ψ. Note that all the plots coincide, demonstrating that r does not depend on ψ, exactly as we assumed in our model. It allows us to express the true positive rate u_ψ as a function of r instead of as a function of the detection threshold. This is shown in Fig. 2. As expected, this curve shows a trade-off between false positive and false negative, and u_ψ is a decreasing function of ψ. We will use the curves in Fig. 2 to provide realistic numbers u_ψ for Section 4. Table 2 lists these numbers.

Table 2: u_ψ *tabulated as a function of r and ψ.*

		ψ				
		2	3	4	5	6
r	0.01	0.83	0.37	0.22	0.12	0.08
	0.05	0.94	0.65	0.42	0.29	0.20
	0.10	0.96	0.75	0.55	0.42	0.31

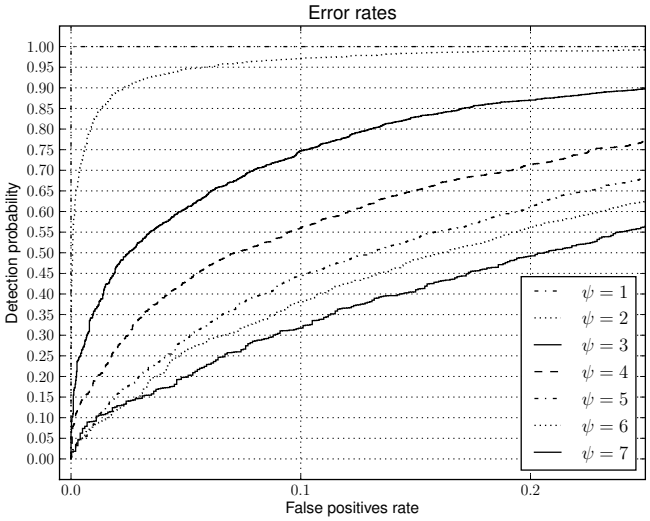

Figure 2: *True positive rate u_ψ as a function of false positive rate r. Shown are the plots for $\psi = 1, \cdots, 7$ (from top to bottom).*

3 Symmetric Tardos fingerprinting code in the combined digit model

For the construction we use a variant of the 'symmetric' Tardos code proposed in [11]. The code generation and embedding steps remain unchanged; only the accusation process is modified to deal with the combined digit model.

3.1 Code generation and embedding

For completeness, we give a brief summary of the code generation and embedding steps, which are a generalization of Tardos's binary code [10]; for more details we refer to [11].

The distributor produces an $n \times m$ matrix $\mathbf{X}$ of q-ary symbols; the rows of the matrix correspond to the fingerprints for the individual users. The matrix is filled in a two-step procedure: The distributor first generates m independent random vectors $\mathbf{p}^{(i)} = (p_0^{(i)}, \cdots, p_{q-1}^{(i)})$ for $1 \leq i \leq m$, where the components satisfy[3] $p_\alpha^{(i)} \in [0, 1]$ and $\sum_{\alpha\in\Sigma} p_\alpha^{(i)} = 1$. We use the notation $\bar{\mathbf{p}} = \{\mathbf{p}^{(i)}\}_{i=1}^m$. The random variables follow a special case of the Dirichlet distribution, $\mathbf{p}^{(i)} \sim F_{q\kappa}$,

$$F_{q\kappa}(\mathbf{p}) = \mathcal{N}_{q\kappa}^{-1} \prod_{\alpha\in\Sigma} p_\alpha^{-1+\kappa} \quad \text{with } \kappa > 0. \tag{1}$$

Here $\mathcal{N}_{q\kappa} = [\Gamma(\kappa)]^q/\Gamma(\kappa q)$ is a normalising constant ensuring that $\int_{J(q)} \mathrm{d}^q\mathbf{p}\ F_{q\kappa}(\mathbf{p}) = 1$. The expression $\int_{J(q)} \mathrm{d}^q\mathbf{p}$ stands for $\int_0^1 \mathrm{d}p_0 \cdots \int_0^1 \mathrm{d}p_{q-1}\delta(1 - \sum_{\beta=0}^{q-1} p_\beta)$, where $\delta(\cdot)$ is the Dirac delta function. The delta function ensures that the integration is done only over $\mathbf{p}$ such that $\sum_\beta p_\beta = 1$. The parameter κ determines the shape of $F_{q\kappa}$. For the binary alphabet one sets $\kappa = 1/2$, reproducing Tardos' distribution function [10].

In the second step, the distributor generates the columns of $\mathbf{X}$ independently. In the i-th column, the vector $\mathbf{p}^{(i)}$ determines the probabilities of generating each specific symbol in the alphabet: $\mathrm{Prob}[X_{ji} = \alpha] = p_\alpha^{(i)}$.

Before the content is released to user j, it is watermarked with the j-th row of the matrix $\mathbf{X}$.

[3]For symplicity we have set the 'cutoff parameter' (see [11]) to zero. This is allowed for nonbinary alphabets.

3.2 Accusation

The distributor extracts the attacked fingerprint y from the unauthorized copy. For each user j, the distributor computes the 'accusation sum' from $\mathbf{X}$, $\bar{\mathbf{p}}$ and y. He decides that the user j is guilty the accusation sum exceeds a threshold Z, where Z is referred to as the 'accusation threshold'. The list of accused users is denoted as $\sigma(\bar{\mathbf{p}}, \mathbf{X}, y)$.

We discuss two possible ways of computing the accusation sum. They both make use of the following weight functions, which were introduced in [10],

$$g(1,p) = \sqrt{\frac{1-p}{p}} \quad ; \quad g(0,p) = -\sqrt{\frac{p}{1-p}} \tag{2}$$

We will often use the notation $g_1(p) = g(1,p)$ and $g_0(p) = g(0,p)$. The weight functions have the special property that

$$pg_1(p) + (1-p)g_0(p) = 0 \quad ; \quad p[g_1(p)]^2 + (1-p)[g_0(p)]^2 = 1. \tag{3}$$

Accusation sum, method A. The watermark detector is applied to y, for every location $1 \leq i \leq m$ for every watermark symbol $\alpha \in \Sigma$ to obtain the values $W_{i\alpha}$. The accusation sum $\mathcal{A}_j$ is computed as

$$\mathcal{A}_j = \sum_{i=1}^{m} \sum_{\alpha \in \Sigma} W_{i\alpha} g(\langle X_{ji} == \alpha \rangle, p_\alpha^{(i)}), \tag{4}$$

where $\langle x \rangle$ returns the value 1 if the Boolean formula x evaluates to TRUE and 0 otherwise. Thus, for each user, m sets of Tardos accusations are summed, scaled by the value $W_{i\alpha}$. The collective accusation sum of the coalition is defined as

$$\mathcal{A}_{\mathrm{C}} = \sum_{j \in C} \mathcal{A}_j = \sum_{i=1}^{m} \sum_{\alpha \in \Sigma} W_{i\alpha} \left\{ b_\alpha^{(i)} g_1(p_\alpha^{(i)}) + [c - b_\alpha^{(i)}] g_0(p_\alpha^{(i)}) \right\}. \tag{5}$$

Here $b_\alpha^{(i)}$ stands for the number of colluders who receive symbol α in segment i.

Accusation sum, method B. We denote with

$$\Phi_i = \{\alpha \in \Sigma : \ W_{i\alpha} = 1\} \tag{6}$$

the set of symbols that are detected at content segment i. The cardinality of this set is $\varphi_i = |\Phi_i|$. We further introduce the notation

$$P_i = \sum_{\alpha \in \Phi_i} p_\alpha^{(i)} \quad ; \quad N_i = \sum_{\alpha \in \Phi_i} b_\alpha^{(i)}. \tag{7}$$

The accusation sum $\mathcal{B}_j$ is computed as

$$\mathcal{B}_j = \sum_{i=1}^{m} g(\langle X_{ji} \in \Phi_i \rangle, P_i). \tag{8}$$

Thus, instead of accusing for each symbol separately as in method A, the symbols are grouped into two sets (detected/undetected in y_i), and a user's accusation is based on the presence of his symbol X_{ji} in one of these sets. The collective accusation sum of the coalition is defined as

$$\mathcal{B}_{\mathrm{C}} = \sum_{j \in C} \mathcal{B}_j = \sum_{i=1}^{m} \left\{ N_i g_1(P_i) + [c - N_i] g_0(P_i) \right\}. \tag{9}$$

Lemma 1 *In the limit of the Unreadable Digit Model ($r = 0$, $u_1 = 1$, $u_\psi = 0$ for $\psi \geq 2$), accusation methods A and B are equivalent.*

Proof: When the coalition embeds more than one symbol into segment i (i.e. $\psi_i \geq 2$), $u_\psi = 0$ causes $W_{i\alpha} = 0$ for all α. Consequently $N_i = 0$ and $P_i = 0$. Eq. (4) then vanishes since $W_{i\alpha} = 0$; Eq. (8) vanishes since $g_0(0) = 0$. When the pirates embed only a single symbol $y_i \in \Sigma$, then $W_{i\alpha} = \delta_{\alpha y}$; Both $\mathcal{A}_j$ and $\mathcal{B}_j$ reduce to $\sum_i g(X_{ji} == y_i, p_{y_i}^{(i)})$. □

4 Analysis

4.1 Symmetry of the attacks

We make two assumptions about the attack strategy ρ. These are the same assumptions as in [11].

1. *Member symmetry*: All members of the coalition are equivalent. The colluders base their decisions only on the number of symbols they receive, and not on the identity of the members who receive them.

2. *Column symmetry*: The strategy for outputting y_i does not explicitly depend on the value i, i.e. the same strategy is used for all y_i. However, we do allow y_i to depend on the full $\mathbf{X}_C$.

The first assumption is motivated by the row symmetry of the code generation and accusation procedures. The second assumption is motivated by the column symmetry of these procedures.

4.2 Performance indicator

The main collusion resistance performance indicator of a fingerprinting code is the coalition size c_0 that can be defeated by a code of a fixed length m, for fixed false positive and false negative error probabilities, for a fixed number of users n. The larger c_0, the better the code.

This can be re-expressed as the code length m required to defeat a coalition of fixed size c_0, for fixed false positive and false negative error probabilities, for a fixed number of users n. The smaller m is, the better the code.

Tardos' binary fingerprinting code [10] achieves

$$m = Gc_0^2\lceil \ln \varepsilon_1^{-1} \rceil, \tag{10}$$

with $G = 100$, and ε_1 the maximum tolerable probability that a fixed innocent user j gets accused (false positive). The false negative (FN) error probability is defined as the probability that none of the colluders get accused. The maximum tolerable FN probability is denoted as ε_2. Tardos set $\varepsilon_2 = \varepsilon_1^{c_0/4}$.

Tardos proved [10] that $m \propto c_0^2$ is asymptotically optimal for any alphabet size. Several works have shown that the parameter G in (10) can be significantly reduced [12, 6, 11, 1] from its original value of 100 by a combination of modifications in the code construction and the proof technique. In particular, in the q-ary code of [11], with ε_2 chosen independently of ε_1 with $\varepsilon_2 \gg \varepsilon_1$, it was shown that the form (10) asymptotically[4] applies for large coalitions, with

$$G = 2\frac{\tilde{\sigma}_{\rm inn}^2}{\tilde{\mu}^2}, \tag{11}$$

where $\tilde{\sigma}_{\rm inn}$ and $\tilde{\mu}$ are statistical parameters of the accusation: $m\tilde{\sigma}_{\rm inn}$ stands for the standard deviation of an innocent user's accusation sum; $m\tilde{\mu}$ stands for the expectation value of the coalition's collective accusation sum. The result (11) holds if an innocent user has zero accusation on average. The symmetric binary scheme of [11] has $\tilde{\sigma}_{\rm inn} = 1$ and $\tilde{\mu} = 2/\pi$, yielding $G = \pi^2/2 \approx 4.9$. Further improvement *in the restricted digit model* is achieved by going to larger alphabets ($q > 2$).

In the coming sections we will use the expression (11) as the main performance indicator of a fingerprinting scheme[5].

4.3 Definition of expectation values

The expectation value taken over all stochastic degrees of freedom will be denoted as $\mathbb{E}$. This includes both stochastic steps of the creation of $\mathbf{X}$, possible randomisation of the colluder strategy ρ (stochastic choice of Ψ_i) and the random behaviour of the inserted noise. We use the notation $\mathbb{E}_p$ for the expectation value over the $\bar{\mathbf{p}}$ degrees of freedom, $\mathbb{E}_X$ for the $\mathbf{X}$ degrees of freedom (at fixed $\bar{\mathbf{p}}$), $\mathbb{E}_\Psi$ for the pirate strategy

[4] The parameter ε_2 appears in terms of relative order $c_0^{-1/2}[\ln \varepsilon_2 / \ln \varepsilon_1]^{1/2}$. These vanish in the regime $c_0 \gg 1$, $\varepsilon_2 \gg \varepsilon_1$.

[5] Under the condition that the expectation value of an innocent user's accusation is zero.

(at fixed $\mathbf{X}_C$) and $\mathbb{E}_W$ for the noise (at fixed Ψ). The full $\mathbb{E}$ can be expressed as $\mathbb{E}_p \circ \mathbb{E}_X \circ \mathbb{E}_\Psi \circ \mathbb{E}_W$. This order reflects the chronological order in which the stochastic events take place. However, other ways of computing $\mathbb{E}$ are possible. In particular, in a number of cases it is convenient to first compute the expectation value over X_j (the j'th row of $\mathbf{X}$, with user j innocent), denoted as $\mathbb{E}_{X_j}$. The $\mathbb{E}_{X_j}$ averaging commutes with the $\mathbf{X}_C$ degrees of freedom and hence with $\mathbb{E}_\Psi$ and $\mathbb{E}_W$.

The $\mathbb{E}_p$ consists of m independent integrals, one for each segment i. Omitting the segment index, we have for each segment:

$$\mathbb{E}_p[\cdots] = \mathcal{N}_{q\kappa}^{-1} \int_{J(q)} \mathrm{d}^q\mathbf{p}\; F_{q\kappa}(\mathbf{p})(\cdots). \tag{12}$$

Likewise, the $\mathbb{E}_{X_C}$ consists of m independent summations over the counting variables $b_\alpha^{(i)}$, one sum per segment. The probability distribution is a multinomial. Omitting the segment index, we have for each segment:

$$\mathbb{E}_{X_C}[\cdots] = \sum_{\vec{b}} \binom{c}{\vec{b}} \prod_{\alpha\in\Sigma} p_\alpha^{b_\alpha}(\cdots) \tag{13}$$

Here it is implicit that $\vec{b}$ satisfies $\sum_{\alpha\in\Sigma} b_\alpha = c$.

4.4 Performance of accusation method A

We first show that we are allowed to use performance indicator (11).

Lemma 2 *In accusation method A, the expectation value of an innocent user's accusation is zero.*

Proof: We compute the expectation value of (4) over X_j, the j'th row of $\mathbf{X}$. We make use of the fact that y (and therefore $W_{i\alpha}$) is independent of X_j when j is innocent.

$$\mathbb{E}_{X_j}[\mathcal{A}_j] = \sum_{i=1}^{m} \sum_{\alpha\in\Sigma} W_{i\alpha} \left[p_\alpha^{(i)} g_1(p_\alpha^{(i)}) + (1-p_\alpha^{(i)}) g_0(p_\alpha^{(i)}) \right]. \tag{14}$$

It follows from the first equation in (3) that the result is zero. □

Before considering arbitrary alphabet size, we first state our result for the binary case.

Theorem 1 *In the case of a binary alphabet ($q = 2$, $\kappa = 1/2$), and assuming $r < \frac{1}{2}$, $u_1 > \frac{1}{2}$, the quantity $\tilde{\sigma}^2_{\mathrm{inn}} := m^{-1}\mathbb{E}[\mathcal{A}_j^2]$ (for innocent j) is upper bounded by $\tilde{\sigma}^2_{\mathrm{inn}} \le [\tilde{\sigma}^{\mathrm{A}}_{\max}]^2$, with*

$$[\tilde{\sigma}^{\mathrm{A}}_{\max}]^2 \le (1-r)u_1 + r(1-u_1) \le 1. \tag{15}$$

Furthermore, independent of the colluder strategy, the expectation value of the collective accusation sum is given by

$$\tilde{\mu}^{\mathrm{A}} = \frac{2}{\pi}(u_1 - r). \tag{16}$$

The performance indicator for method A is upper bounded as

$$G_{\mathrm{A}} \le \frac{\pi^2}{2} \cdot \frac{(1-r)u_1 + r(1-u_1)}{(u_1-r)^2}. \tag{17}$$

Proof: The bound on $\tilde{\sigma}^2_{\mathrm{inn}}$ is proven in Appendix A. The computation of $\tilde{\mu}$ is shown in Appendix B. □

Next we consider non-binary alphabets. We derive a bound on the variance of innocent users' accusation.

Theorem 2 *The quantity $\tilde{\sigma}^2_{\mathrm{inn}} := m^{-1}\mathbb{E}[\mathcal{A}_j^2]$ (for innocent j) is bounded by $\tilde{\sigma}^2_{\mathrm{inn}} \le [\tilde{\sigma}^{\mathrm{A}}_{\max}]^2$, with*

$$[\tilde{\sigma}^{\mathrm{A}}_{\max}]^2 \;:=\; qr + \frac{\Gamma(\kappa q)}{[\Gamma(\kappa)]^q \Gamma(c+\kappa q)} \sum_{\vec{b}} \binom{c}{\vec{b}} \left[\prod_{\alpha\in\Sigma} \Gamma(\kappa + b_\alpha) \right] \max_{\psi\in\{1,\ldots,\omega\}} \psi(u_\psi - r). \tag{18}$$

A proof is given in Appendix A. Note that the theorem does not depend on the colluder strategy.

Corollary 1 *In the limiting case of the unreadable digit model ($r = 0$, $u_1 = 1$, $u_\psi = 0$ for $\psi \geq 2$) Theorem 2 reduces to $[\tilde{\sigma}^{\rm A}_{\rm max}]^2 = 1$.*

Corollary 1 is also proven in Appendix A. The expression (18) looks relatively simple, but direct evaluation of the $\vec{b}$-summation would involve $\mathcal{O}(c^{q-1})$ terms. For numerical evaluation it is more efficient to split the sum up into a sum over ω and sums over the remaining degrees of freedom. After some painful algebra this yields the following result.

Corollary 2 *The bound in Theorem 2 can be rewritten as*

$$[\tilde{\sigma}^{\rm A}_{\rm max}]^2 \;\; = \;\; qr + \frac{c!\,\Gamma(\kappa q)}{\Gamma(c+\kappa q)} \sum_{\omega=1}^{\min(q,c)} \frac{1}{[\Gamma(\kappa)]^\omega} \binom{q}{\omega} S_{c\kappa}(\omega) \max_{\psi\in\{1,\ldots,\omega\}} \psi(u_\psi - r), \tag{19}$$

with

$$S_{c\kappa}(\omega) := \frac{1}{1+\omega(c-\omega)} \sum_{\lambda=0}^{\omega(c-\omega)} e^{i\lambda\frac{2\pi(c-\omega)}{1+\omega(c-\omega)}} \left[\sum_{v=0}^{c-\omega} e^{-i\lambda v\frac{2\pi}{1+\omega(c-\omega)}} \frac{\Gamma(\kappa+1+v)}{(1+v)!}\right]^\omega . \tag{20}$$

The proof is given in Appendix A. While this looks far less transparent than (18), all the summations in (19,20) together require adding only $\mathcal{O}(c^2)$ terms as compared to $\mathcal{O}(c^{q-1})$. For large coalitions (and $q \geq 4$) this can be a significant difference.

Corollary 3 *The variance of an innocent user's accusation satisfies $\tilde{\sigma}^2_{\rm inn} \leq q$.*

Proof: From expression (31) in Appendix A we see that $\tilde{\sigma}^2_{\rm inn}$ is smaller than the expectation value of $\sum_{\alpha\in\Sigma} W_{i\alpha}$ for some arbitrary column index i. As $W_{i\alpha} \in \{0,1\}$ and $|\Sigma| = q$, it follows that $\tilde{\sigma}^2_{\rm inn}$ cannot exceed q. □

Theorem 3 *In accusation method A, it holds for any colluder strategy that $\tilde{\mu} \geq \tilde{\mu}^{\rm A}_{\rm min}$, with*

$$\tilde{\mu}^{\rm A}_{\rm min} := \frac{\Gamma(\kappa q)}{[\Gamma(\kappa)]^q} \frac{c\,c!}{\Gamma(c+\kappa q)} \sum_{\vec{b}} [\prod_{\gamma\in\Sigma} \frac{\Gamma(\kappa+b_\gamma)}{\Gamma(1+b_\gamma)}] \min_{\substack{\Psi\subseteq\Omega\\ \Psi\neq\emptyset}} \left\{ u_\psi \sum_{\alpha\in\Psi} V(b_\alpha) + r \sum_{\alpha\in\Sigma\setminus\Psi} V(b_\alpha) \right\}, \tag{21}$$

where we have defined

$$V(b_\alpha) := \frac{\Gamma(b_\alpha - \frac{1}{2} + \kappa)}{\Gamma(b_\alpha+\kappa)} \frac{\Gamma(c - b_\alpha - \frac{1}{2} + \kappa[q-1])}{\Gamma(c - b_\alpha + \kappa[q-1])} \left\{ \tfrac{1}{2} - \kappa - \frac{b_\alpha}{c}(1-\kappa q) \right\}. \tag{22}$$

A proof is given in Appendix B. Note that the function $V(b_\alpha)$ is exactly the same expression appearing in the Restricted Digit Model treatment in [11]. Setting $r = 0$, $\psi = 1$, $u_1 = 1$ in (21) precisely reproduces the Restricted Digit Model result.

Corollary 4 *In the limit of large c, the quantity $\tilde{\mu}^{\rm A}_{\rm min}$ converges to a finite number.*

A proof is given in Appendix E.

Theorem 4 *The performance parameter G for accusation method A is upper bounded by*

$$G_{\rm A} \leq 2\frac{(\tilde{\sigma}^{\rm A}_{\rm max})^2}{(\tilde{\mu}^{\rm A}_{\rm min})^2}. \tag{23}$$

with $\tilde{\sigma}^{\rm A}_{\rm max}$ and $\tilde{\mu}^{\rm A}_{\rm min}$ as defined in Theorems 2 and 3.

Proof: follows directly from the definition (11) of G and Theorems 2 and 3. □

Unfortunately, this bound is not as sharp as it could be, for two reasons: (i) In the computation of $\tilde{\sigma}^{\mathrm{A}}_{\max}$ we upper-bounded the negative term in (31) by zero, which sacrifices some sharpness. (ii) More importantly, a really sharp bound on the performance parameter would be obtained by a maximization over the colluder strategy: $G_{\mathrm{A}} \leq \max_{\rho}\{2(\tilde{\sigma}^{\mathrm{A}})^2/(\tilde{\mu}^{\mathrm{A}})^2\}$. However, this is a very difficult optimization to carry out, as it amounts to optimally choosing a set Ψ as a function of $\vec{b}$, while the expression $(\tilde{\sigma}^{\mathrm{A}})^2/(\tilde{\mu}^{\mathrm{A}})^2$ depends on Ψ in a very complicated way. We leave this as a subject for future work.

In Fig. 3, G_{A} is plotted for various parameter settings. We see the following trends. For each q, the performance parameter G_A has a minimum as a function of κ, just as in the Restricted Digit Model [11], with almost exactly the same values for the optimal κ. Furthermore, G_{A} increases as a function of r. This is as expected, since the performance should get worse when the attackers become more powerful. A comparison between methods A and B is given in Section 4.6. The results are also compared to the Restricted Digit Model case.

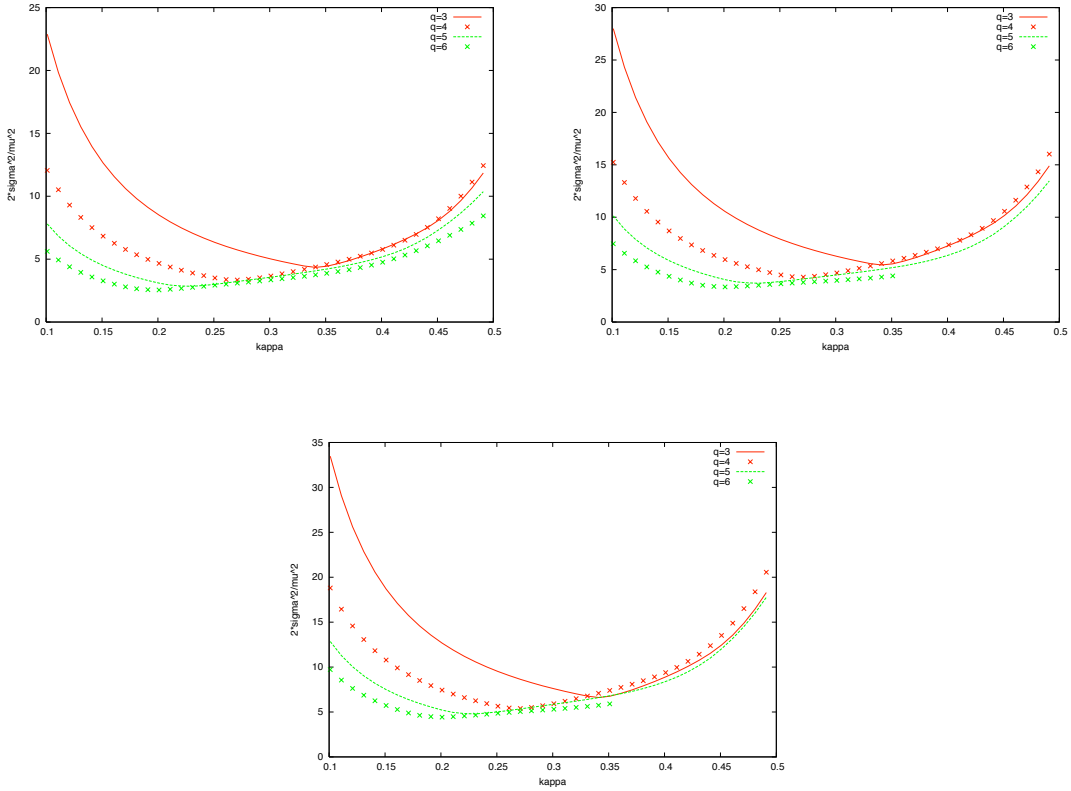

Figure 3: *The bound* $2(\tilde{\sigma}^{\mathrm{A}}_{\max})^2/(\tilde{\mu}^{\mathrm{A}}_{\min})^2$ *on the performance parameter* G *for accusation method A, as a function of* κ*, for* $r = 0.01$*,* $r = 0.05$*,* $r = 0.1$*. In all the graphs we set* $c = 20$*, and* u_ψ *is set according to Table 2.*

4.5 Performance of accusation method B

Lemma 3 *In accusation method B, the expectation value of an innocent user's accusation is zero.*

Proof: For innocent j, we have Prob$[X_{ji} \in \Phi_i] = P_i$. Thus the expectation value of $\mathcal{B}_j$ over X_j is given by

$$\mathbb{E}_{X_j}[\mathcal{B}_j] = \sum_{i=1}^{m} \Big\{ P_i g_1(P_i) + [1-P_i] g_0(P_i) \Big\}. \tag{24}$$

It follows from the first equation in (3) that the result is zero. □

Theorem 5 *The quantity* $(\tilde{\sigma}^{\rm B}_{\rm inn})^2 := m^{-1}\mathbb{E}[\mathcal{B}_j^2]$ *(for an innocent user* j*) is equal to 1.*

Proof: We express $\mathcal{B}_j^2$ as a double sum and note that all the off-diagonal terms disappear (due to column independence) when the expectation $\mathbb{E}_{X_j}$ is taken. Finally the second equation in (3) is used.

$$\begin{aligned}
\frac{1}{m}\mathbb{E}_{X_j}[\mathcal{B}_j^2] &= \frac{1}{m}\sum_{i,k=1}^{m} \mathbb{E}_{X_j}\Big[g([X_{ji}\in\Phi_i], P_i)\; g([X_{jk}\in\Phi_k], P_k)\Big] \\
&= \frac{1}{m}\sum_{i=1}^{m} \mathbb{E}_{X_j}\left[g^2([X_{ji}\in\Phi_i], P_i)\right] + \frac{1}{m}\sum_{i,k:\; i\neq k} 0 \\
&= \frac{1}{m}\sum_{i=1}^{m} \big\{ P_i g_1^2(P_i) + (1-P_i) g_0^2(P_i)\big\} = 1.
\end{aligned}$$

□

Theorem 6 *In accusation method B, for the binary alphabet, it holds for any colluder strategy that*

$$\tilde{\mu}^{\rm B} = \frac{2}{\pi}(u_1 - r). \tag{25}$$

The performance parameter is given by

$$G_{\rm B} = \frac{\pi^2}{2}\cdot\frac{1}{(u_1-r)^2}. \tag{26}$$

The proof is given in Appendix C.

Theorem 7 *In accusation method B, it holds for any colluder strategy that* $\tilde{\mu} \geq \tilde{\mu}^{\rm B}_{\rm min}$*, with*

$$\begin{aligned}
\tilde{\mu}^{\rm B}_{\rm min} \;:=\; & (1-r)^q \frac{c!\,\Gamma(\kappa q)}{\Gamma(c+\kappa q)} \sum_{\omega=1}^{\min(c,q)} \frac{1}{[\Gamma(\kappa)]^\omega}\binom{q}{\omega} \sum_{\substack{\vec{v}\in\{0,\ldots,c-\omega\}^\omega \\ \sum_k v_k = c-\omega}} \left[\prod_{a=1}^{\omega}\frac{\Gamma(\kappa+1+v_a)}{\Gamma(2+v_a)}\right] \\
& \min_{\substack{\lambda\in\{0,1\}^\omega \\ |\lambda|\neq 0}} \sum_{\zeta\in\{0,1\}^\omega} \left(\frac{u_{|\lambda|}}{r}\right)^{\lambda\cdot\zeta} \left(\frac{1-u_{|\lambda|}}{1-r}\right)^{\lambda\cdot(\lambda-\zeta)} \sum_{x=0}^{q-\omega}\binom{q-\omega}{x}[\frac{r}{1-r}]^\varphi \\
& \Big\{\frac{c}{2} - c\kappa\varphi - N + N\kappa q\Big\} \frac{\Gamma(-\frac{1}{2}+N+\kappa\varphi)}{\Gamma(N+\kappa\varphi)} \frac{\Gamma(-\frac{1}{2}+c-N+\kappa[q-\varphi])}{\Gamma(c-N+\kappa[q-\varphi])}
\end{aligned} \tag{27}$$

where $\varphi = x + |\zeta|$ *and* $N = |\zeta| + \zeta\cdot v$*. The* $|\;|$ *notation stands for the Hamming weight, and the* $\cdot$ *is the inner product.*

A proof is given in Appendix C. Unfortunately expression (27) is not very transparent. The main reason is that taking the expectation value $\mathbb{E}_W$ is rather involved, as we have to keep track of both used and unused symbols in Ω, which have different detection probabilities.

Corollary 5 *In the limit of large* c *the quantity* $\tilde{\mu}^{\rm B}_{\rm min}$ *converges to a finite number.*

The proof is given in Appendix F.

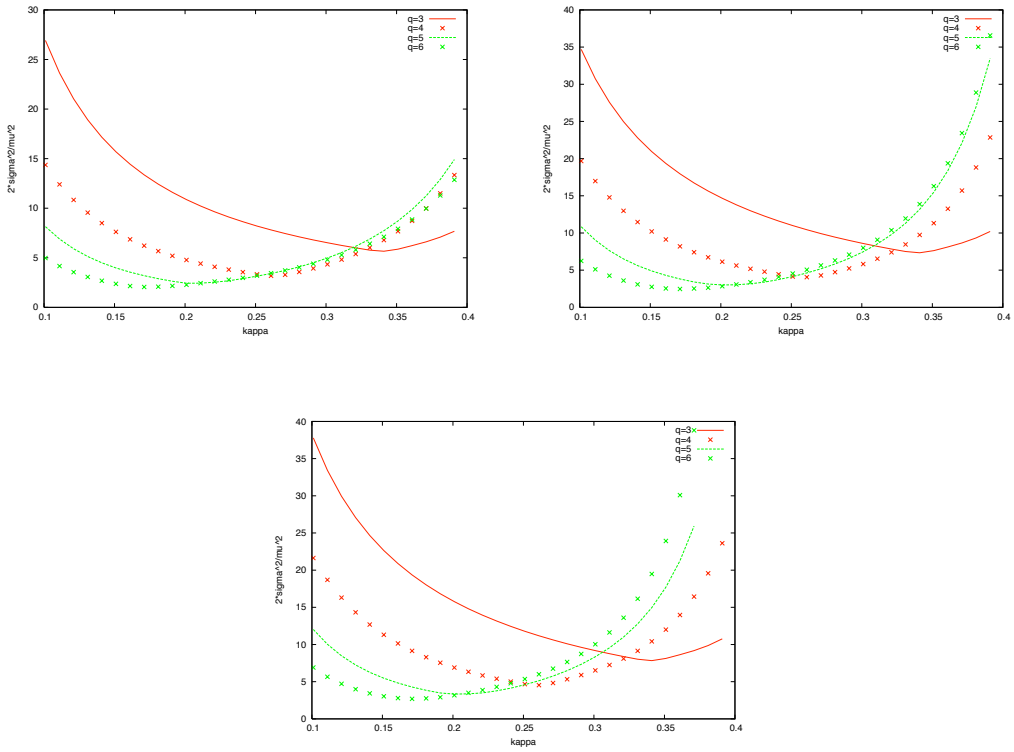

Figure 4: *The bound* $2/(\tilde{\mu}^{\mathrm{B}}_{\min})^2$ *on the performance parameter* G *for accusation method B, as a function of* κ*, for* $r = 0.01$*,* $r = 0.05$*,* $r = 0.1$*. In all the graphs we set* $c = 20$*, and* u_ψ *is set according to Table 2.*

Theorem 8 *For accusation method B, the performance parameter* G *is upper bounded by*

$$G_{\mathrm{B}} \leq 2/(\tilde{\mu}^{\mathrm{B}}_{\min})^2, \tag{28}$$

with $\tilde{\mu}^{\mathrm{B}}_{\min}$ *as defined in Theorem 7.*

Proof: follows directly from the definition (11) of G, Theorem 5 and Theorem 7. □

In contrast to method A, this bound is sharp. Furthermore it directly points at a 'worst case' pirate strategy that forces the content owner to use a low code rate. This is precisely the strategy that minimizes $\tilde{\mu}$ by choosing (for each combination $\{\omega, \vec{v}\}$ separately, i.e. for each $\vec{b}$) the string λ (which is equivalent to the set Ψ) such that the expression after $\min_\lambda$ in (27) is minimized. Clearly this is not a trivial strategy.

In Fig. 4, G_{B} is plotted for various parameter settings. We see the same trends as in method A. For each q, the performance parameter G_B has a minimum as a function of κ, just as in the Restricted Digit Model [11], with almost exactly the same values for the optimal κ. Furthermore, G_{B} increases as a function of r, as expected. A comparison between methods A and B is given in Section 4.6. The results are also compared to the Restricted Digit Model case.

4.6 Comparison

The numerical results are summarized in Table 3. For each of the curves in Figs. 3 and 4 we have taken the minimum, and listed the optimal κ and G value in the table. We have also included the results from [11] for the Restricted Digit Model. It is clear that methods A and B do not differ dramatically. (That was also

Table 3: *The performance parameter G for accusation methods A and B in the Combined Digit Model, and for the Restricted Digit Model. The listed κ is the optimal value for given q, for fixed $c = 20$.*

		Method A		Method B		Restricted
q	r	κ	G_{A}	κ	G_{B}	Digit Model
3	0.01	0.34	4.2	0.34	5.6	κ=0.34
	0.05	0.34	5.4	0.34	7.3	G=2.6
	0.1	0.34	6.6	0.34	7.8	
4	0.01	0.27	3.3	0.26	3.1	κ=0.26
	0.05	0.27	4.2	0.26	4.0	G=1.9
	0.1	0.27	5.5	0.26	4.5	
5	0.01	0.23	2.8	0.21	2.5	κ=0.23
	0.05	0.23	3.7	0.21	3.0	G=1.6
	0.1	0.22	4.8	0.20	3.3	
6	0.01	0.20	2.5	0.17	2.1	κ=0.19
	0.05	0.20	3.5	0.17	2.5	G=1.4
	0.1	0.20	4.5	0.17	2.8	

the case in [13], where the code was studied for a different attack model, and with a different performance indicator.) Method A is better at $q = 3$, and method B is better[6] at $q \geq 4$.

What is most striking is that even a strong attack ($r = 0.1$) does not seriously reduce the effectiveness of the code. Compared to the Restricted Digit Model, the code length has to be increased by less than a factor 2.5. We conclude that accusation methods A and B are quite effective for dealing with the increased attack strength in the Combined Digit Model.

5 Summary

We have introduced a new attack model for coalition attacks on watermarks, the Combined Digit Model. The model comprises averaging attacks and signal processing attacks in a way that is more realistic than the Unreadable Digit Model. The detector may detect multiple symbols in the same content segment. The probability of false positive detection is represented as a single parameter r. The false negative error probabilities are represented as a vector u_ψ, where ψ is the number of symbols mixed together by the colluders. The r and u_ψ parameters all depend on the detection threshold. However, since r is almost uniquely determined by the detection threshold, independent of ψ, it is possible to think of the vector u_ψ as being a function of r.

We have examined two modifications of the accusation sum in the symbol-symmetric Tardos scheme. Method A sums up $g_{0/1}$ accusations for each detected symbol separately. Method B groups all detected symbols together and then applies the accusation function $g_{0/1}$. We have evaluated the performance of both schemes in terms of the performance parameter $G := 2\tilde{\sigma}^2_{\mathrm{inn}}/\tilde{\mu}^2$ which is based on the Gaussian approximation as introduced in [12]. For the binary alphabet the results are very simple, and it turns out that method A is slightly better.

For nonbinary alphabets we have obtained analytic expressions for G. These unfortunately do not look very insightful, but they do enable efficient numerical evaluation. It turns out that methods A and B have similar performance. Method B is better at $q \geq 4$. The q-ary Tardos code with either of the modified accusation methods is effective against powerful attacks in the Combined Digit Model.

[6] Reminder: The numbers for method A are perhaps pessimistic, as mentioned in Section 4.4.

References

[1] O. Blayer and T. Tassa. Improved versions of Tardos' fingerprinting scheme. *Designs, Codes and Cryptography*, 48(1):79–103, 2008.

[2] D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. *IEEE Transactions on Information Theory*, 44(5):1897–1905, 1998.

[3] T. Furon, A. Guyader, and F. Cérou. On the design and optimization of Tardos probabilistic fingerprinting codes. In *Information Hiding*, Lecture Notes in Computer Science, pages 341–356. Springer, 2008.

[4] H.-J. Guth and B. Pfitzmann. Error- and collusion-secure fingerprinting for digital data. In *Information Hiding*, volume 1768 of *Springer Lecture Notes in Computer Science*, pages 134–145. Springer, 1999.

[5] S. He and M. Wu. Joint coding and embedding techniques for multimedia fingerprinting. 1:231–248, June 2006.

[6] K. Nuida, M. Hagiwara, H. Watanabe, and H. Imai. Optimal probabilistic fingerprinting codes using optimal finite random variables related to numerical quadrature. *CoRR*, abs/cs/0610036, 2006.

[7] H.G. Schaathun. Novel attacks on spread-spectrum fingerprinting. *EURASIP Journal of Information Security*, page Article ID 803217, 2008. Open access at `http://www.hindawi.com/getarticle.aspx?doi=10.1155/2008/803217&e=ref`.

[8] H.G. Schaathun. On error-correcting fingerprinting codes for use with watermarking. *Multimedia Systems*, 13(5-6):331–344, 2008.

[9] M. Steinebach, J. Dittmann, and E. Saar. Combined fingerprinting attacks against digital audio watermarking: methods, results and solutions. In B. Jerman-Blažič and T. Klobučar, editors, *Communications and Multimedia Security*, volume 228 of *IFIP Conference Proceedings*, pages 197–212. Kluwer, 2002.

[10] G. Tardos. Optimal probabilistic fingerprint codes. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC)*, pages 116–125, 2003.

[11] B. Škorić, S. Katzenbeisser, and M.U. Celik. Symmetric Tardos fingerprinting codes for arbitrary alphabet sizes. *Designs, Codes and Cryptography*, 46(2):137–166, 2008.

[12] B. Škorić, T.U. Vladimirova, M.U. Celik, and J.C. Talstra. Tardos fingerprinting is better than we thought. *IEEE Transactions on Information Theory*, 54(8):3663–3676, 2008.

[13] F. Xie, T. Furon, and C. Fontaine. On-off keying modulation and Tardos fingerprinting. In A.D. Ker, J. Dittmann, and J.J. Fridrich, editors, *MM&Sec*, pages 101–106. ACM, 2008.

[14] Hong Zhao, Min Wu, Z. June Wang, and K. J. Ray Liu. Nonlinear collusion attacks on independent fingerprints for multimedia. *IEEE Trans. Image Proc.*, pages 646–661, 2005.

A Evaluation of $\tilde{\sigma}_{\rm inn}$ for method A

We start by bounding the expression $\mathbb{E}_{X_j}[\mathcal{A}_j^2]$ for an innocent user j. Note that $W_{i\alpha}$ and X_{ji} ($j \notin C$) are independent.

$$\begin{aligned}
\mathbb{E}_{X_j}[\mathcal{A}_j^2] &= \sum_{i,k=1}^{m} \sum_{\alpha,\beta\in\Sigma} W_{i\alpha}W_{k\beta}\, \mathbb{E}_{X_j}\left[g(\langle X_{ji}=\alpha\rangle, p_\alpha^{(i)})g(\langle X_{jk}=\beta\rangle, p_\beta^{(k)})\right] \\
&= \sum_{i=1}^{m} \sum_{\alpha,\beta\in\Sigma} W_{i\alpha}W_{i\beta}\, \mathbb{E}_{X_j}\left[g(\langle X_{ji}=\alpha\rangle, p_\alpha^{(i)})g(\langle X_{ji}=\beta\rangle, p_\beta^{(i)})\right] \qquad (29)
\end{aligned}$$

Here we have used the fact that all off-diagonal terms ($k \neq i$) vanish due to the first property in (3). Next we split the double sum $\sum_{\alpha\beta}$ into terms with $\beta = \alpha$ and terms with $\beta \neq \alpha$.

$$\begin{aligned}\mathbb{E}_{X_j}[\mathcal{A}_j^2] &= \sum_{i=1}^{m}\sum_{\alpha\in\Sigma} W_{i\alpha}^2 + \sum_{i=1}^{m}\sum_{\substack{\alpha,\beta\in\Sigma\\ \alpha\neq\beta}} W_{i\alpha}W_{i\beta}\Big[p_\alpha^{(i)}g_1(p_\alpha^{(i)})g_0(p_\beta^{(i)}) + p_\beta^{(i)}g_0(p_\alpha^{(i)})g_1(p_\beta^{(i)}) \\ &\quad + (1-p_\alpha^{(i)}-p_\beta^{(i)})g_0(p_\alpha^{(i)})g_0(p_\beta^{(i)})\Big] \end{aligned} \tag{30}$$

Again using the first property in (3) we simplify this to

$$\mathbb{E}_{X_j}[\mathcal{A}_j^2] = \sum_{i=1}^{m}\sum_{\alpha\in\Sigma} W_{i\alpha}^2 - \sum_{i=1}^{m}\sum_{\substack{\alpha,\beta\in\Sigma\\ \alpha\neq\beta}} W_{i\alpha}W_{i\beta}g_0(p_\alpha^{(i)})g_0(p_\beta^{(i)}). \tag{31}$$

Binary alphabet

In the case of a binary alphabet, (31) reduces to

$$\mathbb{E}_{X_j}[\mathcal{A}_j^2] = \sum_{i=1}^{m}(W_{i1}-W_{i0})^2. \tag{32}$$

Note that $(W_{i1}-W_{i0})^2 \in \{0,1\}$. When $\psi = 1$ we have $\mathbb{E}_W[(W_{i1}-W_{i0})^2] = (1-r)u_1 + r(1-u_1)$, whereas for $\psi = 2$ we have $\mathbb{E}_W[(W_{i1}-W_{i0})^2] = 2u_2(1-u_2)$. This allows us to write the following strategy-independent bound,

$$\mathbb{E}_W[(W_{i1}-W_{i0})^2] \leq \max\left\{2u_2(1-u_2),\ (1-r)u_1 + r(1-u_1)\right\}. \tag{33}$$

Assuming that $r < 1/2$ and $u_1 > 1/2$, the second expression in the 'max' is always larger than the first. The result (33) is a constant, so the expectation values $\mathbb{E}_\Psi$, $\mathbb{E}_X$ and $\mathbb{E}_p$ are trivial. This immediately leads to the result given in Theorem 1.

Non-binary alphabet

We bound[7] (31) as $\mathbb{E}_{X_j}[\mathcal{A}_j^2] \leq \sum_{i=1}^{m}\sum_{\alpha\in\Sigma} W_{i\alpha}$. Assuming column symmetry of the attack (see Section 4.1), we get $\tilde{\sigma}_{\rm inn} \leq \mathbb{E}_p\mathbb{E}_{X_C}\mathbb{E}_y\mathbb{E}_W\sum_{\alpha\in\Sigma} W_{i\alpha}$.

Next, applying $\mathbb{E}_W$ to $W_{i\alpha}$ gives $\mathbb{E}_W[W_{i\alpha}] = u_{\psi_i}$ if $\alpha \in \Psi_i$ and r if $\alpha \notin \Psi_i$. Thus,

$$\sum_{\alpha\in\Sigma}\mathbb{E}_W[W_{i\alpha}] = (q-\psi_i)r + \psi_i u_{\psi_i}. \tag{34}$$

Note that this expression depends on the set Ψ_i only through the integer $\psi_i \leq \omega_i$. Next we apply $\mathbb{E}_\Psi$. We bound the result as

$$\mathbb{E}_\Psi\mathbb{E}_W\sum_{\alpha\in\Sigma} W_{i\alpha} \leq qr + \max_{\psi_i\in\{1,\ldots,\omega_i\}}\psi_i(u_{\psi_i}-r). \tag{35}$$

We briefly remark on the limiting case ($r = 0, u_1 = 1, u_\psi = 0$ for $\psi \geq 2$) corresponding to the *unreadable digit model*. In this limit the $\alpha \neq \beta$ terms in (31) vanish. Furthermore (35) is trivially upper bounded by 1, yielding $[\tilde{\sigma}_{\rm inn}^{\rm A}]^2 \leq 1$.

We return to the Combined Digit Model. Note that (35) depends only on ω_i. It is independent of the colluder strategy and independent of all other columns $\neq i$. Hence, in applying $\mathbb{E}_{X_C}$ and $\mathbb{E}_p$ we only have to deal with column i. From here on we drop the column index i.

$$\begin{aligned}\tilde{\sigma}_{\rm inn}^2 &\leq \frac{1}{m}\mathbb{E}_p\mathbb{E}_{X_C}\mathbb{E}_\Psi\mathbb{E}_W\sum_{i=1}^{m}\sum_{\alpha\in\Sigma} W_{i\alpha} \\ &\leq qr + \sum_{\vec{b}}\binom{c}{\vec{b}}\left(\mathbb{E}_p\prod_{\alpha\in\Sigma}p_\alpha^{b_\alpha}\right)\max_{\psi\in\{1,\ldots,\omega\}}\psi(u_\psi - r)\end{aligned} \tag{36}$$

[7] When ω is small, the colluders' safest choice is to embed a single symbol. In that case the product $W_{i\alpha}W_{i\beta}$ is zero with high probability. On the other hand, when ω is large, then the colluders do a powerful averaging attack yielding a small $\mathbb{E}_W[W_{i\alpha}]$. In that case the product $\mathbb{E}_W[W_{i\alpha}W_{i\beta}]$ is much smaller than $\mathbb{E}_W[W_{i\alpha}]$. Furthermore, the product $g_0(p_\alpha)g_0(p_\beta)$ cannot exceed 1, as $p_\alpha \leq 1-p_\beta$ and $p_\beta \leq 1-p_\alpha$. Hence we expect our bound to be reasonably sharp.

Note that ω is a function of $\vec{b}$. We use the following well known property of Dirichlet integrals,

$$\int_0^1 \mathrm{d}^q p\ \delta(1-\sum_{\alpha\in\Sigma} p_\alpha) \prod_{\alpha\in\Sigma} p_\alpha^{-1+x_\alpha} = \frac{\prod_{\alpha\in\Sigma}\Gamma(x_\alpha)}{\Gamma(\sum_{\alpha\in\Sigma} x_\alpha)} \tag{37}$$

to obtain

$$\int_{J(q)} \mathrm{d}^q\mathbf{p}\ F_{q\kappa}(\mathbf{p}) \prod_{\alpha\in\Sigma} p_\alpha^{b_\alpha} = \frac{\Gamma(\kappa q)}{[\Gamma(\kappa)]^q} \frac{\prod_{\alpha\in\Sigma}\Gamma(\kappa+b_\alpha)}{\Gamma(c+\kappa q)}. \tag{38}$$

Substitution of (38) into (36) gives (18). That completes the proof of Theorem 2. For Corollary 2 we have to further evaluate the sum $\sum_{\vec{b}}$. We make use of the fact that the summand is fully symbol-symmetric, i.e. it is invariant under any permutation of the alphabet. This allows us to split $\sum_{\vec{b}}$ into a sum over ω (with combinatorial multiplicity $\binom{q}{\omega}$) times a sum over the leftover counting variables $v_1,\cdots,v_\omega$ which keep track of how the leftover $c-\omega$ colluders are divided over the ω symbols in Ω. We have $v_k = b_{\alpha_k} - 1$ with $\alpha_k \in \Omega$, and $\sum_{k=1}^{\omega} v_k = c-\omega$.

$$\sum_{\vec{b}} \binom{c}{\vec{b}} \to \sum_{\omega=1}^{\min(q,c)} \binom{q}{\omega} \sum_{\vec{v}} \frac{c!}{\prod_{k=1}^{\omega}(1+v_k)!} \tag{39}$$

In this representation, the expression (18) becomes

$$\tilde{\sigma}^2_{\mathrm{inn}} \quad\leq\quad qr + \frac{c!\,\Gamma(\kappa q)}{\Gamma(c+\kappa q)} \sum_{\omega=1}^{\min(c,q)} [\Gamma(\kappa)]^{-\omega} \binom{q}{\omega} S_{c\kappa}(\omega) \max_{\psi\in\{1,\ldots,\omega\}} \psi(u_\psi - r), \tag{40}$$

$$S_{c\kappa}(\omega) \quad:=\quad \sum_{\vec{v}:\ \sum_{s=1}^{\omega} v_s = c-\omega} \prod_{k=1}^{\omega} \frac{\Gamma(\kappa+1+v_k)}{(1+v_k)!}. \tag{41}$$

We further evaluate $S_{c\kappa}(\omega)$ by rewriting the constrained $\vec{v}$-sum as an unconstrained sum with a Kronecker delta in the summand.

$$S_{c\kappa}(\omega) \quad=\quad \sum_{v_1=0}^{c-\omega} \cdots \sum_{v_\omega=0}^{c-\omega} \delta_{c-\omega,\sum_k v_k} \frac{\Gamma(\kappa+1+v_1)}{(1+v_1)!} \cdots \frac{\Gamma(\kappa+1+v_\omega)}{(1+v_\omega)!}. \tag{42}$$

Finally we use a sum representation of the Kronecker delta,

$$\delta_{ab} = \frac{1}{M} \sum_{\lambda=0}^{M-1} e^{i\lambda(a-b)\frac{2\pi}{M}}, \tag{43}$$

with $M = \omega(c-\omega)+1$, to obtain a factorisation of the sums $\sum_{v_k}$. The expression for $S_{c\kappa}(\omega)$ in (20) follows. □

B Evaluation of $\tilde{\mu}$ for method A

We define

$$Q_\alpha(\vec{b}^{(i)}) := \mathbb{E}_{\bar{\mathbf{P}}\setminus\mathbf{p}^{(i)}} \mathbb{E}_{X_C\setminus X_C^{(i)}} \mathbb{E}_\Psi \mathbb{E}_W [W_{i\alpha}]. \tag{44}$$

Following steps that are completely analogous to the analysis in [11], we arrive at

$$\tilde{\mu} = \frac{\Gamma(\kappa q)}{[\Gamma(\kappa)]^q} \frac{c\ c!}{\Gamma(c+\kappa q)} \sum_{\vec{b}} [\prod_{\gamma=0}^{q-1} \frac{\Gamma(\kappa+b_\gamma)}{\Gamma(1+b_\gamma)}] \sum_{\alpha\in\Sigma} Q_\alpha(\vec{b}) V(b_\alpha), \tag{45}$$

with $V(b_\alpha)$ as defined in (22). Here we have dropped the column index i because of the column symmetry. For given $\vec{b}$ and non-empty $\Psi \subseteq \Omega$ we can write

$$\sum_{\alpha\in\Sigma} V(b_\alpha)\mathbb{E}_W[W_{i\alpha}] = u_\psi \sum_{\alpha\in\Psi} V(b_\alpha) + r \sum_{\alpha\in\Sigma\setminus\Psi} V(b_\alpha) \geq \min_{\substack{\Lambda\subseteq\Omega \\ \Lambda\neq\emptyset}} \left\{ u_{|\Lambda|} \sum_{\alpha\in\Lambda} V(b_\alpha) + r \sum_{\alpha\in\Sigma\setminus\Lambda} V(b_\alpha) \right\}. \tag{46}$$

The last expression is independent of the colluder strategy. Hence, application of the expectation values $\mathbb{E}_{\bar{\mathbf{p}}\setminus\mathbf{p}^{(i)}}\mathbb{E}_{X_C\setminus X_C^{(i)}}\mathbb{E}_\Psi$ leaves the expression unchanged. Substitution into (45) yields the result given in Theorem 3.

Binary alphabet

When the alphabet is binary we have $q = 2$ and $\kappa = 1/2$. (This value of κ reproduces Tardos' distribution function for p, namely $f(p) \propto [p(1-p)]^{-1/2}$.) Taking the limit $\kappa \to 1/2$ in (22) is almost trivial. The factor $\{\cdots\} = (\frac{1}{2} - \kappa)(1 - 2b_\alpha/c)$, which goes to 0, makes $V(b_\alpha)$ vanish, except for one subtlety. When $b_\alpha = 0$ or $b_\alpha = c$, one of the Gamma functions in the numerator goes to $\Gamma(0) = \infty$. Using $\lim_{\kappa\to 1/2}(\kappa - 1/2)\Gamma(\kappa - 1/2) = 1$ we get

$$V(c) = -V(0) = \frac{\Gamma(c)}{\Gamma(1/2)\Gamma(c + 1/2)}, \tag{47}$$

and $V(b_\alpha) = 0$ for $b_\alpha \in \{1, \cdots, c-1\}$. Hence there are only two surviving terms in the $\vec{b}$-sum in (45): all ones and all zeroes. In the former case $\Psi = \{1\}$, $\alpha \in \Psi$ implies $b_\alpha = c$, and $\alpha \notin \Psi$ implies $b_\alpha = 0$. In the latter case, $\Psi = \{0\}$, $\alpha \in \Psi$ implies $b_\alpha = 0$, and $\alpha \notin \Psi$ implies $b_\alpha = c$. Substituting (47) into (45) in this way yields $\tilde{\mu} = \frac{2}{\pi}(u_1 - r)$.

C Evaluation of $\tilde{\mu}$ for method B

We start from the collective accusation sum (9) and take the expectation value, making use of column symmetry.

$$\tilde{\mu} = \mathbb{E}_p\mathbb{E}_X\mathbb{E}_\Psi\mathbb{E}_W[N_i g_1(P_i) + (c - N_i)g_0(P_i)]. \tag{48}$$

Here the column index i is arbitrary, and we will omit it when this does not cause ambiguity.

Binary alphabet

For $q = 2$ computing the expectation values is very simple. We note that the expression $Ng_1(P) + [c - N]g_0(P)$ vanishes for $(W_0, W_1) = (1, 1)$ and $(W_0, W_1) = (0, 0)$. In the former case because $N = c$, $P = 1$, yielding $c - N = 0$ and $g_1(P) = 0$; In the latter case because $N = 0$, $P = 0$, yielding $g_0(P) = 0$. This leaves only the combinations $(W_0, W_1) = (0, 1)$ and $(W_0, W_1) = (1, 0)$ to evaluate. We note that the $(0, 1)$ case gives $Ng_1(P) + [c - N]g_0(P) = xg_1(p_1) + [c - x]g_0(p_1)$, while the $(1, 0)$ case gives $Ng_1(P) + [c - N]g_0(P) = -[xg_1(p_1) + [c - x]g_0(p_1)]$. Here x denotes the number of ones received by the coalition. It follows that

$$\begin{aligned}&\mathbb{E}_W[Ng_1(P) + [c - N]g_0(P)] = \\ &\quad \{xg_1(p_1) + [c - x]g_0(p_1)\} \left(\Pr[W_0 = 0, W_1 = 1] - \Pr[W_0 = 1, W_1 = 0] \right).\end{aligned} \tag{49}$$

The expectation value $\mathbb{E}_X$ can be written as $\mathbb{E}_X[\cdots] = \sum_{x=0}^{c} \binom{c}{x} p_1^x (1-p_1)^{c-x}(\cdots)$. It was shown in [10] that

$$\mathbb{E}_p[\{xg_1(p_1) + [c - x]g_0(p_1)\}p_1^x(1 - p_1)^{c-x}] = \frac{1}{\pi}(\delta_{x,c} - \delta_{x,0}). \tag{50}$$

(Here Tardos' cutoff parameter t has been set to zero.) Only the terms $x = 0$ and $x = c$ in $\mathbb{E}_X$ survive. For $x = c$, it follows from the marking condition that $\Psi = \{1\}$. Similarly, for $x = 0$ it follows that $\Psi = \{0\}$.

Hence (48) evaluates to

$$\begin{aligned}\tilde{\mu} &= \frac{1}{\pi}\left(\Pr[W_0=0,W_1=1|\Psi=\{1\}]-\Pr[W_0=1,W_1=0|\Psi=\{1\}]\right)\\ &\quad -\frac{1}{\pi}\left(\Pr[W_0=0,W_1=1|\Psi=\{0\}]-\Pr[W_0=1,W_1=0|\Psi=\{0\}]\right).\end{aligned} \tag{51}$$

The probabilities in (51) are given by $\Pr[W_0=1,W_1=0|\Psi=\{0\}]=\Pr[W_0=0,W_1=1|\Psi=\{1\}]=(1-r)u_1$ and $\Pr[W_0=1,W_1=0|\Psi=\{1\}]=\Pr[W_0=0,W_1=1|\Psi=\{0\}]=r(1-u_1)$, resulting in $\tilde{\mu}=\frac{2}{\pi}(u_1-r)$.

Non-binary alphabet

The case $q\geq 3$ is not nearly as simple as the binary case. Again we start from (48). Since N_i and P_i depend only on the colluders' degrees of freedom in X, $\mathbb{E}_X$ is equivalent to $\mathbb{E}_{X_C}$. Substituting (13) and (12) into (48), we can write

$$\begin{aligned}\tilde{\mu} &= \sum_{\vec{b}^{(i)}}\binom{c}{\vec{b}^{(i)}}\mathbb{E}_{\vec{p}\setminus p^{(i)}}\mathbb{E}_{X_C\setminus X_C^{(i)}}\mathbb{E}_\Psi\mathbb{E}_W\\ &\quad\left[N_i\mathbb{E}_{p^{(i)}}[g_1(P_i)\prod_{\alpha\in\Sigma}[p_\alpha^{(i)}]^{b_\alpha^{(i)}}]+(c-N_i)\mathbb{E}_{p^{(i)}}[g_0(P_i)\prod_{\beta\in\Sigma}[p_\beta^{(i)}]^{b_\beta^{(i)}}]\right].\end{aligned} \tag{52}$$

Both $\mathbb{E}_p$ integrals can be evaluated exactly. The method is shown in Appendix D. The result is

$$\begin{aligned}\mathbb{E}_p[g_1(P)\prod_{\alpha\in\Sigma}p_\alpha^{b_\alpha}] &= \mathcal{N}_{q\kappa}^{-1}\frac{\Gamma(-\frac{1}{2}+\kappa\varphi+N)\Gamma(\frac{1}{2}+\kappa(q-\varphi)+c-N)}{\Gamma(c+\kappa q)}\\ &\quad\times\frac{\prod_{\alpha\in\Sigma}\Gamma(\kappa+b_\alpha)}{\Gamma(N+\kappa\varphi)\Gamma(c-N+\kappa[q-\varphi])},\end{aligned} \tag{53}$$

$$\begin{aligned}\mathbb{E}_p[g_0(P)\prod_{\alpha\in\Sigma}p_\alpha^{b_\alpha}] &= -\mathcal{N}_{q\kappa}^{-1}\frac{\Gamma(\frac{1}{2}+\kappa\varphi+N)\Gamma(-\frac{1}{2}+\kappa(q-\varphi)+c-N)}{\Gamma(c+\kappa q)}\\ &\quad\times\frac{\prod_{\alpha\in\Sigma}\Gamma(\kappa+b_\alpha)}{\Gamma(N+\kappa\varphi)\Gamma(c-N+\kappa[q-\varphi])}.\end{aligned} \tag{54}$$

Here we have omitted the segment index i on P_i, N_i, φ_i, $b_\alpha^{(i)}$, and $p_\alpha^{(i)}$. We bound $\tilde{\mu}$ using

$$\mathbb{E}_\Psi[\cdots]\geq\min_{\Psi\subseteq\Omega:\ \Psi\neq\emptyset}(\cdots). \tag{55}$$

This gets rid of any strategy dependence. Consequently, the operations $\mathbb{E}_{\vec{p}\setminus p^{(i)}}$ and $\mathbb{E}_{X_C\setminus X_C^{(i)}}$ have no effect on the bound. Next we reorganize the sum $\sum_{\vec{b}}$ as in (39). In this way we obtain a bound $\tilde{\mu}\geq\tilde{\mu}_{\min}^B$, with

$$\begin{aligned}\tilde{\mu}_{\min}^B &= \frac{\Gamma(\kappa q)\ c!}{\Gamma(c+\kappa q)}\sum_{\omega=1}^{\min(c,q)}\frac{1}{[\Gamma(\kappa)]^\omega}\binom{q}{\omega}\sum_{\vec{v}}[\prod_{k=1}^{\omega}\frac{\Gamma(\kappa+1+v_k)}{\Gamma(2+v_k)}]\min_{\substack{\Psi\subseteq\{1,\ldots,\omega\}\\ \Psi\neq\emptyset}}\mathbb{E}_W\\ &\quad\left\{\frac{c}{2}-c\kappa\varphi-N+N\kappa q\right\}\frac{\Gamma(-\frac{1}{2}+\kappa\varphi+N)\Gamma(-\frac{1}{2}+\kappa[q-\varphi]+c-N)}{\Gamma(N+\kappa\varphi)\Gamma(c-N+\kappa[q-\varphi])}.\end{aligned} \tag{56}$$

Finally we write the expectation $\mathbb{E}_W$ as a double sum: one over symbols in Ω and another over symbols $\notin\Omega$. We represent Ψ as a string $\lambda\in\{0,1\}^\omega$, with $\lambda_b=1$ if $\alpha_b\in\Psi$. We represent W as a combination of an integer $x\in\{0,\ldots,q-\omega\}$ and a string $\zeta\in\{0,1\}^\omega$. The x counts the number of detected symbols that the colluders did not have at their disposal. $\zeta_b=1$ indicates that the b'th symbol in Ω is detected. Combined with the detection probability, this gives

$$\mathbb{E}_W[\cdots]=\sum_{x=0}^{q-\omega}\binom{q-\omega}{x}\sum_{\zeta\in\{0,1\}^\omega}u_{|\lambda|}^{\zeta\cdot\lambda}(1-u_{|\lambda|})^{|\lambda|-\zeta\cdot\lambda}r^{x+|\zeta|-\zeta\cdot\lambda}(1-r)^{q-|\lambda|-x-|\zeta|+\zeta\cdot\lambda}(\cdots), \tag{57}$$

where the notation $|\lambda|$ stands for the Hamming weight of λ, and $\zeta \cdot \lambda$ is the inner product $\sum_{b=1}^{\omega} \zeta_b \lambda_b$. The quantities φ and N are expressed as $\varphi = x + |\zeta|$ and $N = |\zeta| + \zeta \cdot v$. Substitution of (57) into (56) yields (27). □

D Evaluation of the integrals in Appendix C

The integrals (53) and (54) are evaluated as follows. We have to compute a q-dimensional integral of the form

$$\mathcal{N}_{q\kappa}^{-1} \int_0^1 \mathrm{d}^q p \; \delta(1 - \sum_{\gamma \in \Sigma} p_\gamma) [\prod_{\alpha \in \Sigma} p_\alpha^{-1+\kappa}] [\prod_{\alpha \in \Sigma} p_\alpha^{b_\alpha}] g_{0/1}(\sum_{\beta \in \Phi} p_\beta). \tag{58}$$

We split the q-dimensional integration space into a part belonging to Φ and a part outside Φ. For $\alpha \in \Phi$ we write $p_\alpha = P s_\alpha$, and for $\beta \notin \Phi$ we write $p_\beta = (1-P) t_\beta$. The integration splits as

$$\begin{aligned}
&\int_0^1 \mathrm{d}^q p \; \delta(1 - \sum_{\alpha \in \Sigma} p_\alpha) = \int_0^1 \mathrm{d}^q p \; \delta(1 - \sum_{\alpha \in \Sigma} p_\alpha) \int_0^1 \mathrm{d}P \; \delta(P - \sum_{\beta \in \Phi} p_\beta) \\
&= \int_0^1 \mathrm{d}P \; P^\varphi (1-P)^{q-\varphi} \int_0^P \mathrm{d}^\varphi s \int_0^{1-P} \mathrm{d}^{q-\varphi} t \; \delta(P - P \sum_{\alpha \in \Phi} s_\alpha) \delta(1 - P - [1-P] \sum_{\beta \in \Sigma \setminus \Phi} t_\beta) \\
&= \int_0^1 \mathrm{d}P \; P^{\varphi-1} (1-P)^{q-\varphi-1} \int_0^1 \mathrm{d}^\varphi s \; \delta(1 - \sum_{\alpha \in \Phi} s_\alpha) \int_0^1 \mathrm{d}^{q-\varphi} t \; \delta(1 - \sum_{\beta \in \Sigma \setminus \Phi} t_\beta).
\end{aligned} \tag{59}$$

Furthermore, the two products appearing in the integrand can be split in the same way:

$$\begin{aligned}
\prod_{\alpha \in \Sigma} p_\alpha^{b_\alpha} &= P^N (1-P)^{c-N} \prod_{\alpha \in \Phi} s_\alpha^{b_\alpha} \prod_{\beta \in \Sigma \setminus \Phi} t_\beta^{b_\beta}. \\
\prod_{\alpha \in \Sigma} p_\alpha^{-1+\kappa} &= P^{\varphi(-1+\kappa)} (1-P)^{(q-\varphi)(-1+\kappa)} \prod_{\alpha \in \Phi} s_\alpha^{-1+\kappa} \prod_{\beta \in \Sigma \setminus \Phi} t_\beta^{-1+\kappa}.
\end{aligned} \tag{60}$$

With this split, each full q-dimensional integral gets factorized into three independent integrals. The φ-dimensional s-integral and the $(q-\varphi)$-dimensional t-integral are evaluated using (37). The one-dimensional P-integral yields a Beta function. Multiplying the pieces together yields (53) and (54).

E Convergence of $\tilde{\mu}_{\min}^{\mathrm{A}}$ in the large-c limit

In this appendix we look at $\tilde{\mu}_{\min}^{\mathrm{A}}$ in the limiting case where c becomes very large. The following lemma helps us in determining the asymptotic behaviour of gamma functions.

Lemma 4 *For $x \gg 1$ and constants a, b with $|a| \ll x$ and $|b| \ll x$, it holds that*

$$\frac{\Gamma(x+a)}{\Gamma(x+b)} = x^{a-b}[1 + \mathcal{O}(1/x)].$$

Proof: Follows directly from Stirling's approximation. □

Using Lemma 4, we see that $V(b_\alpha)$ in (22) scales as c^{-1}, the product $\prod_\gamma$ in (21) scales as $c^{q(\kappa-1)}$ and the quotient $\frac{c!}{\Gamma(c+\kappa q)}$ as $c^{1-q\kappa}$. Furthermore, the number of terms in $\sum_{\vec{b}}$ scales as c^{q-1}. (The '-1' comes from the constraint $\sum_\alpha b_\alpha = c$, which reduces the number of degrees of freedom by one.) Combining all the powers of c contained in (21) in this way, we find c^0. Hence $\tilde{\mu}_{\min}^{\mathrm{A}}$ converges.

F Convergence of $\tilde{\mu}^{\mathrm{B}}_{\min}$ in the large-c limit

For large c, the N and v_a scale as c^1. From Lemma 4 in Appendix E it follows that the x-sum in (27) scales as c^0, the product $\prod_a \frac{\Gamma(\cdots)}{\Gamma(\cdots)}$ scales as $c^{\omega(\kappa-1)}$ and the quotient $\frac{c!}{\Gamma(c+\kappa q)}$ as $c^{1-\kappa q}$. The $\vec{v}$-summation has $\mathcal{O}(c^{\omega-1})$ terms. Using $\omega \leq q$ and combining all the powers of c we get c^0 as the highest power of c occurring in (27).