Plausibilistic Entropy and Anonymity*

Iulian Goriac†
Department of Computer Science
"Al.I.Cuza" University
Iasi, Romania
iulian.goriac@info.uaic.ro

Abstract

A common approach behind measuring anonymity is that the larger the anonymity set is the higher the degree of anonymity it supports. Our approach builds upon this intuition proposing a very general and yet precise measure for security properties. Introduced in a paper accepted for ARES 2013 conference, plausibilistic entropy promises to offer an expressive and cost effective solution for quantifying anonymity. This article focuses on a detailed side-by-side comparison between plausibilistic entropy and Shannon entropy and underlines a promising level of compatibility between the two of them. Towards the end we present our vision on how to define a measure for anonymity based on plausibilistic entropy and how such a definition can be employed to serve practical purposes.

Keywords: plausibilistic entropy, knowledge, multi-agent systems, anonymity

1 Introduction

Because of the interest the scientific community showed in the newly introduced concept of plausibilistic entropy by accepting the publication of the article "Measuring Anonymity with Plausibilistic Entropy" at the ARES conference [1], we would like now to present a more thorough analysis of this concept by rolling out a side-by-side comparison with the classical probabilistic entropy.

What this article aims to provide is a solution to the problem of quantifying security properties like anonymity. In most practical implementations anonymity is presented as dilution of the perceived responsibility of a certain agent in relation to a certain action. The larger the set of agents that could have performed the action the greater the degree of anonymity provided by the setting. We believe that the reasoning behind this approach is sound, and, it is our belief that this is a base to be built upon. Still, taking into consideration only the size of the anonymity set is not enough to guarantee what Halpern and O'Neill call "the divorce between the actions and the agents who perform them for some set of observers". This happens because, in some cases, the very action that we try to separate from the agent can carry information capable of identifying it. What we are looking for is a formula that can factor in all this meta-information about the system, calculate a reliable degree of anonymity, and do that in a cost effective manner. For example, let us say that an observer learns that one of the members of an anonymity group wears a 'red sweater' and this makes it less likely than another specific member to have performed a certain action (the relation being applicable for the two agents only). Using probabilities could be a way of expressing this but this introduces at least two problems. The first one is related to the actual probability values to be assigned to the two agents: there are no statistics about "red sweaters" available! The second one: even if we could assign values, how do we express the fact that a certain agent

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, volume: 5, number: 1, pp. 64-83

*This paper is an extended version of the work originally presented at the 8th International Conference on Availability, Reliability and Security (ARES'13), Regensburg, Germany, September 2013 [1].

†Corresponding author: Tel: +40-771-248-476

is less likely than another to have performed the action while not introducing implicitly any assumption regarding the relation between any of the two agents and the rest of the group?

This is why we consider a plausibility based solution. Not only can the plausibilistic approach make abstraction of the actual numeric values but it can also be used to express more reliably the relations between the likelihoods. The problem with the plausibilistic spaces is that they are very scarce in properties and this makes it difficult if one tries to compare them. The concept of plausibilistic entropy that this article introduces is aimed at making this problem easier to manage.

For the concerned reader, this paper also rounds up some threads started in the previous work of the authors of [2, 3, 1] whose purpose is to compile a unifying epistemic logic based framework for expressing and (automatic) reasoning about information security concepts.

Previous work: For defining security related properties, the closest sources that the formalism employed here can be traced to are the works of Halpern and O'Neill [4, 5, 6, 7] on anonymity (possibilistic and probabilistic approaches) and epistemic logics. Deng et al. [8] hint to the opportunity of using entropy to measure/characterise anonymity. The credit for the idea of using a general plausibilistic approach for defining security properties goes to Halpern in [9, 10, 11]. The concept of entropy introduced by Shannon [12] to measure the degree of uncertainty in a system was very influential for defining plausibilistic entropy as a corresponding measure applicable for plausibility spaces. In the ARES paper [1] the concept of plausibilistic entropy was introduced and the definition for plausibilistic α group anonymity. For graph rendering and symbolic calculation we used the Sage mathematics software [13].

Contributions: In addition to the ARES 2013 published article, this paper comes to provide an in-depth analysis of the plausibilistic entropy by comparing its basic properties to those of the classical probabilistic entropy. The results summarised in Table 1 indicate a promising degree of compatibility between the two notions. Abstracting the details behind the definition of plausibilistic entropy, which can still be found in [1], this material indexes the definitions of anonymity that are compatible with our formal framework and imagines a practical application for the new definition of anonymity using the Herbivore protocol [14] as a case study.

Document structure: Section 2 briefly introduces the fundamental notions we use to discuss about security properties. This includes definitions for MAS, protocols, runs, knowledge, and an epistemic logic. Section 3 extensively presents plausibilistic entropy and compares it to the classical probabilistic entropy. A brief discussion related to conditioning plausibilities is also reproduced. Section 4 collects a small set of definitions for anonymity that are compatible with our framework and envisions a practical application for the plausibilistic definition of anonymity.

2 Fundamentals

In this section we introduce the basic notions that will help us describe the 'actors' of our presentation (the *agents*) and their interaction. Formal notions and terminology are based on [4, 5, 6].

2.1 Entities

Definition 2.1 (MAS). *A* multi-agent system *is a tuple* $\mathscr{S} = (Ag, G, Act)$ *where:*

$Ag = \{A_1, A_2, \ldots, A_n\}$ *is a finite non-empty set of* agents*, capable of storing and processing information. All the information an agent A has access to at a certain moment in time is encapsulated in its* local state*,* $l_A \in L_A$*, where* L_A *is the set of all possible local states the agent A can have. Sometimes it is useful to consider the environment, E, as a special element in Ag;*

$G = L_{A_1} \times L_{A_2} \times \cdots \times L_{A_n}$ *is the set of all the* global states *a system can have. Any element* $g = (l_{A_1}, l_{A_2}, \ldots, l_{A_n}) \in G$ *is called a global state and tuples the local states of all the agents in Ag;*

$Act = \{a_1, a_2, \ldots\}$ *is a set of* actions*. Actions are initiated by agents and are defined by the change they introduce in the global state of the system.* Act_A *denotes the set of actions agent A can be associated to.*

2.2 Protocols and runs

The MASs that we are dealing with are not static, they change from moment to moment as a consequence of the modifications that agents introduce by their actions. We propose a formalisation of this based on the concept of protocol from [4]. A detailed elaboration of these concepts can be found in [3].

Definition 2.2 (protocol)**.** *A* protocol *for an agent A in a system* $\mathscr{S}$ *is a function* $P_A : L_A \to 2^{Act_A}$ *specifying the set of actions agent A can perform in every local state. A tuple* $\mathscr{P} = (P_{A_1}, P_{A_2} \ldots P_{A_n})$ *describing the actions that can be performed by each agent in an MAS is called a* joint protocol*.*

An MAS with a corresponding (joint) protocol can be described as a transition system with the states represented by the elements in G and the transition arcs labelled by action vectors in $Act_{A_1} \times Act_{A_2} \times \cdots \times Act_{A_n}$.

Definition 2.3 (run)**.** *A* run *is a function* $r : T \to G$ *from a set we call time, T, to the set of global states of an MAS.*

Definition 2.4 (point)**.** *If r is a run and m is a moment in time,* $m \in T$*, then the pairing* (r,m) *is called a* point*. Every point has a global state associated to it,* $r(m) = g \in G$*, and by* $r_x(m) = l_{A_x} \in L_{A_x}$ *is denoted the* x^{th} *component of the associated global state* $r(m)$*. If R is a set of runs then* $\wp(R)$ *is used to indicate the set of all points in R.*

Remark: R is normally used to denote a certain non-empty set of runs. In the continuation of the material R will implicitly specify all the possible runs of the MAS.

2.3 Knowledge of an agent

If we consider all the points of an MAS we realise that we cannot always bijectively map all the global states of an MAS to the local states of any single agent (unless we consider the system as being (or consisting of) only one agent). Take for instance an agent A and a set of runs R. It is possible for agent A to have the same local state in at least two different global states. Because A only has access to its local state then it will be impossible for it to distinguish between the previously mentioned global states, and therefore the agent might 'think' that any of them is possible in that moment. Sometimes it can be useful to view a run as the *context* in which the agent operates then realise that the agent might not always be fully aware of what is happening around it.

Definition 2.5 (agent-information set)**.** *Given* $\mathscr{S}$ *an MAS with a protocol* $\mathscr{P}$*, a set of runs R, an agent A, and a point* (r,m)*, the set of all points in* $\wp(R)$ *that A thinks are possible at* (r,m) *is defined as*

$$\mathscr{K}_A(r,m) = \left\{(r',m') \in \wp(R) \mid r'_A(m') = r_A(m)\right\}$$

and we call it an agent-information set *[5].*

Definition 2.6 (indistinguishability)**.** *Two points are* indistinguishable *w.r.t. an agent A if both of them belong to the same agent-information set* $\mathscr{K}_A(r,m)$*. This is denoted by* $(r,m) \sim_A (r',m')$*.*

To conclude this section, we say that, intuitively, an agent A *knows* a fact φ at a point (r,m) if φ is true at all points in $\mathscr{K}_A(r,m)$ [6]. This intuition will now be formalised.

2.4 An epistemic logic

To provide support for reasoning about various security related properties we employ the precise semantics in [6]. The syntax of the epistemic logic is recursively defined by using a set Φ of primitive propositions, two Boolean logic operators ($\neg$ and $\wedge$), and a modal operator K_A representing the knowledge of an agent A. The semantics is defined via the concept of *interpreted system*.

Definition 2.7 (interpreted system)**.** *An* interpreted system *is defined as a tuple $\mathscr{I} = (\mathscr{S}, \mathscr{P}, R, \pi)$ with $\mathscr{S}$, $\mathscr{P}$ and R having the usual interpretation and π being a point dependent interpretation assigning truth values to all the primitive propositions $p \in \Phi$: $(\pi(r,m))(p) \in \{true, false\}$.*

Definition 2.8 (truth)**.** *Given an interpreted system $\mathscr{I}$, an agent A, and a group of agents $G \subseteq Ag$, the* truth value *of a formula φ at point (r,m) is recursively defined by:*

$$
\begin{array}{lll}
(\mathscr{I}, r, m) \models p & \text{iff} & (\pi(r,m)(p)) = true; \\
(\mathscr{I}, r, m) \models \neg\varphi & \text{iff} & (\mathscr{I}, r, m) \not\models \varphi; \\
(\mathscr{I}, r, m) \models \varphi \wedge \psi & \text{iff} & (\mathscr{I}, r, m) \models \varphi \text{ and } (\mathscr{I}, r, m) \models \psi; \\
(\mathscr{I}, r, m) \models K_A[\varphi] & \text{iff} & (\mathscr{I}, r', m') \models \varphi \text{ for all } (r', m') \in \mathscr{K}_A(r, m);
\end{array}
$$

Additionally, the dual operator of $K_A[\varphi]$, meaning that "A knows the fact φ", is defined as $P_A[\varphi] = \neg K_A[\neg\varphi]$ with the intuitive meaning of "A thinks that φ is possible". For convenience, the following notations will also be used from now on:

$\mathscr{I} \models \varphi$**:** meaning that φ is *valid* in $\mathscr{I}$, or formally $(\mathscr{I}, r, m) \models \varphi$ for all $(r,m) \in \mathscr{I}$;

$\forall_{(X)}$**:** *conjunction* over the formulas in set X;

$\theta(A,a)$**:** a primitive proposition — agent A *has performed* or *will perform* action a during the same run;

I**:** will identify a special agent called *observer*.

3 Plausibilistic entropy

In Section 2.3 we were discussing about the local state of an agent not always being able to help the agent distinguish between the various global states the surrounding system might be in. We also defined knowledge as a very special condition in which the agent can be certain that a fact is valid (when that very fact is actually true in all the global states containing the same specific local state). No indication was given so far as to how an agent could learn that it actually knows something. If we were to switch our point of view from that of the overseer of the system to that of a regular agent, the first thing we would become aware of is the fact that, given only the information of the local state, we could not always tell for *certain* in what state the system is, meaning that we could not tell the validity of a certain fact contained in our local information storage. With a bit of luck[1] we might however get a hint about which global states are possible or, even better, we could exhaustively enumerate all the states that are possible given the information that we have (the assumption behind this is that the changes in the system somehow consistently affect our local storage). An even better situation is met when we can introduce a structure over the set of possible global states (or "worlds") that we think are possible (the meaning of the possibility operator $P_A[\varphi]$ can be applied here). We are now making the step from the possibilistic approach to epistemology (where we are either certain that a certain fact is true or we will not take it into

[1] "luck" here stands for an appropriate experience that allegedly provided us with the chance to properly map external states to local ones (the reality strategy)

consideration) to a plausibilistic approach (where knowledge, as defined here, is just an intangible ideal case and decisions are usually made based on uncertain information). Intuitively, given a certain set of possible worlds, the *uncertainty* will be maximum when we will consider each of them equally *likely* and minimum we will have decided that strictly one applies[2]. The plausibilistic approach allows us to consider any point along this interval not just its extremes. Let us formalise this a now...

Definition 3.1 (plausibility space [1])**.** *A* plausibility space *is a triple* $(\Omega, \mathscr{F}, \nu)$ *with:*

Ω *is a set of possible outcomes (e.g. all the worlds an agent thinks that are possible at a certain point, all the runs of an MAS, all global states, etc.);*

$\mathscr{F}$ *is a* σ*-algebra over* Ω *(containing* $\emptyset$ *and closed to complementation and countable union) — when* $\mathscr{F}$ *is not explicitly specified we will consider it to be* 2^{Ω}*;*

ν *is a* plausibility measure *mapping any element in* $\mathscr{F}$ *to some arbitrary set D (partially ordered by the* $\leq$ *relation and containing two special elements* $\bot$ *and* $\top$ *with* $\bot \leq d \leq \top$, $\forall d \in D$*) —* ν *has the following properties:*

1. $\nu(\emptyset) = \bot$*;*
2. $\nu(\Omega) = \top$*;*
3. *if* $\Omega_1, \Omega_2 \subseteq \mathscr{F}$ *with* $\Omega_1 \subseteq \Omega_2$ *then* $\nu(\Omega_1) \leq \nu(\Omega_2)$.

The plausibility spaces and plausibility measures are very general approaches to representing uncertainty, the structure of D being able to store various kinds of information an agent has about the possible surrounding worlds. For instance if D contains only one element d (other then $\bot$ and $\top$) the agent can 'feel' absolutely certain about its surroundings. Intuitively, the less 'organised' the structure of D the less certain the agent is about what is happening outside its local state. Imagine A having more than one option ($|D| > 3$) and not being able to tell which one of them is preferable (or more plausible) and compare this to D being totally ordered. Imagine D varying in size and ramification, capable of supporting various operations (like addition and multiplication), and so on... This becomes important when A has to make a decision, when A has to choose between the possible worlds in order to take the most appropriate action. Ideally we should have a unique measure for A's level of uncertainty allowing us to define security properties not only in a possibilistic/qualitative manner but in a plausibilistic/quantitative way as well. Compare the following situations: the observer I, not knowing for certain either of the facts φ, ψ, or χ, thinks that φ is 'more possible' than ψ, or that the 'distance' between the likelihood of φ and that of ψ is greater than the one between the likelihood of ψ and that of χ or even that φ is ten times more 'likely' than χ. For the special case of a *probability structure*[3], such measures were defined and are commonly referred to as *entropy* ([15] compiles a list of such approaches).

In [12], Shannon defines a measure for entropy over a discrete set of possible events whose probabilities of occurrence are $p_1, p_2, \ldots, p_n$ to be a function $H(p_1, p_2, \ldots, p_n)$ with the following properties:

1. H is continuous in p_i.
2. If all the p_i are equal, $p_i = \frac{1}{n}$, then H is a monotonic increasing function of n. With equally likely events there is more choice, or uncertainty, when there are more possible events.
3. If a choice [can] be broken down into two successive choices, the original H should be the weighted sum of the individual values of H.

[2] decision making can be viewed as an (inapplicable) worlds elimination process

[3] $D = [0,1]$, $\bot = 0$, $\top = 1$ and the third property of ν is replaced with the following:
if $\Omega_1, \Omega_2 \subseteq \mathscr{F}$ *with* $\Omega_1 \cap \Omega_2 = \emptyset$ *then* $\nu(\Omega_1 \cup \Omega_2) = \nu(\Omega_1) + \nu(\Omega_2)$

Afterwards, a theorem is proven showing that the only H satisfying all of the three assumptions above looks like:

$$H = -K\sum_{i=1}^{n} p_i \log(p_i)$$

where K is a positive constant.

3.1 Definition

The notion that we are about to introduce here will help us make quantitative statements about the degree of certainty/uncertainty an agent has related to a fact.

Our question is: how can we define an entropy measure over a plausibility space? In order to answer it, we will try an approach, similar to the one Shannon used to define the entropy of a probability space: first, some intuitive properties will be investigated and then a formal expression will be attempted. Because our ultimate goal is to quantify group anonymity given a plausibility structure, we admit that our intuitions can be biased toward achieving this purpose, however, we do not deny the fact that this approach might be useful in other contexts too. Our approach to entropy comes from a decision making perspective:

1. Any two plausibility structures must be comparable, regardless of their sizes and/or complexities. What we want to know is whether a certain plausibility structure is preferable to another in terms of how well it protects an agent from the attacks of an intruder. It is easier to make a rational choice/decision when you have a set of values or some sort of ordering.

2. If all n elements of a plausibility structure (except $\top$ and $\bot$ of course) have the same plausibility (are in the same layer) then the entropy should be a monotonic increasing function of n. The similarity with the second requirement in the probabilistic approach is obvious. If an observer cannot tell whether an agent is more likely to have taken a certain action than any other agent then the larger the group of suspects the more difficult to find the culprit.

3. The entropy of a plausibility structure represents the amount of uncertainty in that structure, the more variants at any level the greater the entropy.

Figure 1 shows some examples of plausibility structures that will be referenced in our explanations. Let us consider first that each of the structures there indicates the plausibility structure of some corresponding anonymity set, each set containing exactly five agents: A_1 to A_5 (for simplicity, no agents are mapped to $\top$ or $\bot$). The mapping is made top down and left to right, higher positions indicate higher plausibilities. The connecting edges are actually directed edges, pointing downwards. Our guideline is: if you wanted to guess the agent that did the action in which situation would you rather be?

Intuitively, according to the aforementioned guidelines the most desirable situation is D_{12}, when the plausibility structure is totally ordered. Notice that, in this case, regardless of the level we are at, there is no choice to be made, entitling us to consider this a minimal or no entropy situation (at least for sets of finite size). *The number of choices available for a certain node in a plausibility structure equals the number of emerging edges minus one ($\bot$ is not taken into consideration).* Thus the amount of choice for D_{12} literally equals 0.

At the opposite end of the spectrum we have D_{11}, indicating the fact that there is no way of discriminating between agents, giving us the greatest amount of choice: 4. The plausibilistic entropy is definitely directly proportional with the volume of choice in a structure. By this simple measure the entropies for the structures in Figure 1 would be: $4-0-1-1$, $3-3-3$, and $3-5-3$ respectively for structures aligned in the three rows of the figure.

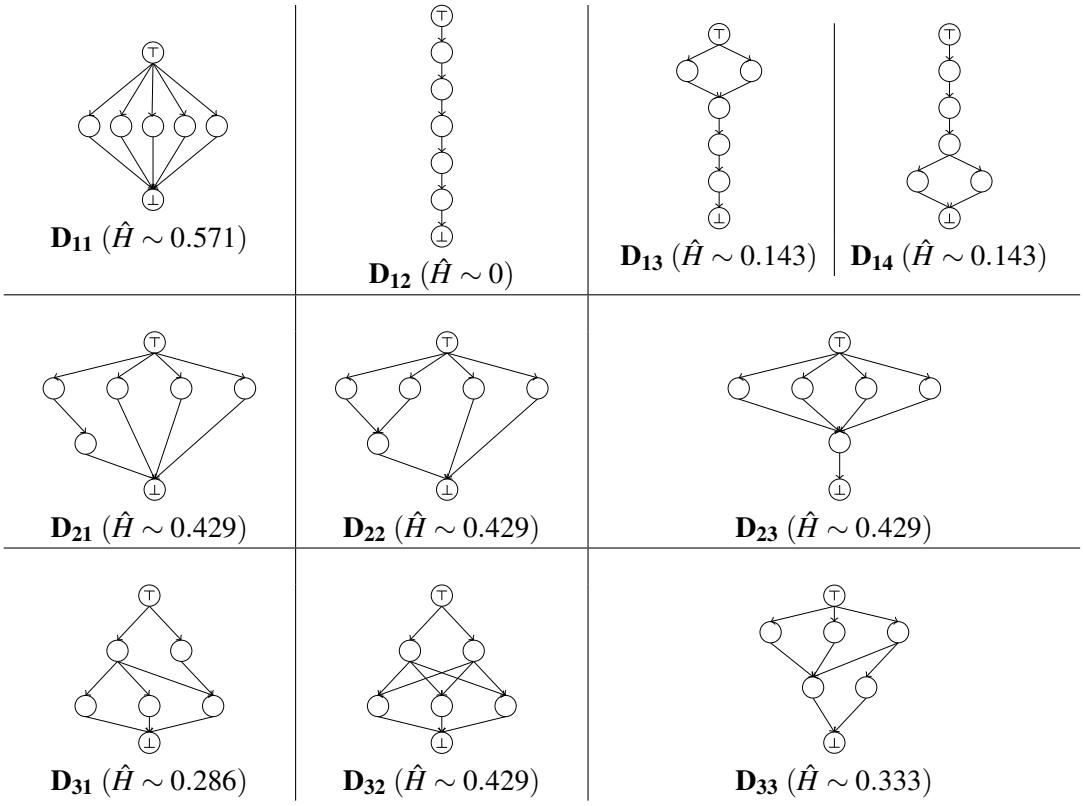


Figure 1: Some plausibility structures for $|D| = 7$ (this figure details Fig. 2 in [1])

This would render the entropy of D_{32} greater than that of D_{11}. Yet, from a decision making perspective, situation D_{32} is preferable to D_{11} because it justifies the choice between the agents in the top layer over the ones in the bottom layer. The two points of view can be reconciled if we consider that *the contribution of a choice node to the systemic entropy depends on the structure of the layer that particular node is contained into*, more specifically *the number of nodes equally distanced from* $\top$. There is obviously more choice in layer 2 of D_{32} than in the same layer of D_{31}.

Another thing to consider is the number of elements in a plausibility structure. Initially we would be tempted to state that the larger the set the larger the entropy. However, adding a new node will not necessarily add more choice: consider D_{23} over a flat 4 agent anonymity set – namely D_4 – (D_{11} can be characterised as a flat 5 agent anonymity set); both would have the same amount of choice (layered or not). Despite its more complex structure D_{23} can provide more certainty than D_4 because the newly added node does not alter the original hierarchy but rather supports it, like a new fact that does not contradict an existing theory. If this sounds a little counter-intuitive we have to realise the fact that the larger the structure the more facts it can deal with and knowledge is what ultimately reduces the entropy of the structure by ordering it. Generally, the more an observer learns the better prepared it is for making a decision. Even if the new knowledge might initially increase the entropy of the structure in the long run the effort might pay off. Therefore we will consider that *plausibilistic entropy (again from a decision making point of view) is inversely proportional to the number of elements in the plausibility set*.

Remark: For simplicity, because plausibility structures are in fact directed acyclic graphs, in the following demonstrations we will use some graph theory notions from [16].

Definition 3.2 (normalised plausibility structure [1])**.** *A plausibility structure D is* normalised *if, given two elements $v_1, v_2 \in D$ with $v_1 \leq v_2$ and $\exists v \in D$ with $v_1 \leq v$, and $v \leq v_2$, then $v = v_1$ or $v = v_2$. For this*

definition to make sense the transitivity property of $<$ be artificially and temporarily suppressed.

Definition 3.3 (distance [1])**.** *Given a normalised plausibility structure D, the* distance *between an element $v \in D$ and the top element $\top$, $dist(\top, v)$ or $dist_\top(v)$, is the number of edges on the longest path from $\top$ to v. Consequently the distance between two elements $v_1, v_2 \in D$ is $dist(v_1, v_2) = |dist_\top(v_1) - dist_\top(v_2)|$ with $|\cdot|$ representing the absolute value.*

Definition 3.4 (layer [1])**.** *Given a normalised plausibility structure D, by* layer *we understand the set of all the elements in D equally distanced from $\top$. The layers are 0 based indexed. Thus the top layer has index 0, $L_0 = \{\top\}$, and for every $k > 0$ $L_k = \{v \mid dist_\top(v) = k\}$, $v \in D$.*

As a consequence of the previous discussion, we propose the following formula for quantifying the entropy of a plausibility structure:

Definition 3.5 (plausibilistic entropy [1])**.** *Given D a normalised plausibility space/structure (as used in Definition 3.1) the* plausibilistic entropy *of D is the sum of the average amounts of choice per layer divided by the number of elements in D:*

$$\hat{H} = \frac{\sum_{k=0}^{l-2} \left(\frac{\sum_{v \in L_k} d_D^+(v)}{|L_k|} - 1 \right)}{n}$$

where:

l is the number of layers in D ($\top$ and $\bot$ containing layers included);

L_k is the set of elements in layer k;

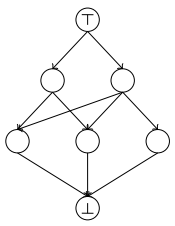
$d_D^+(v)$ is the number of edges emerging from v;

n is $|D|$, the number of elements in D.

In this case $|\cdot|$ represents the number of elements in a set.

Proposition 3.1. *Plausibilistic entropy is well defined.*

Proof. By definition D must have at least two elements, $\bot$ and $\top$ making $n \geq 2$, so the denominator of the main fraction will never be 0. Since $\bot \leq \top$, D will always have at least 2 layers so the first sum of the main numerator will always make sense. For the top fraction, any layer must have at least one element (otherwise it would not be a layer at all) thus making $|L_i| > 0$ so its denominator will never be 0 either. In conclusion our formula is well defined mathematically for any plausibility structure. □



$$\begin{array}{lcccl} \hat{h}_0 & = & \frac{(2-1)}{1} & = & 1 \\ \hat{h}_1 & = & \frac{(2-1)+(3-1)}{2} & = & 1.5 \\ \hat{h}_2 & = & \frac{(1-1)+(1-1)+(1-1)}{3} & = & 0 \\ \hat{H} & = & \frac{1+1.5+0}{7} & \sim & 0.375\ldots \end{array}$$

Figure 2: A step by step example for calculating the plausibilistic entropy of a structure.

A complementary formula for plausibilistic entropy can be taken into consideration (layers are indexed relative to $\perp$):

$$\check{H} = \frac{\sum_{k=0}^{l-2}\left(\frac{\sum_{v\in L_k} d_D^-(v)}{|L_k|} - 1\right)}{n}$$

and to make the distinction between the two formulas they will be referred by *top plausibilistic entropy* and respectively *bottom plausibilistic entropy*.

Remark: The choice between the two variant depends on whether the context is related to finding the highest plausibility element or the lowest. For instance, in the case of *anonymity* we will use $\hat{H}$, while for *onymity* [17] $\check{H}$ might turn out to be a more appropriate alternative.

3.2 Properties

Proposition 3.2. *Given a family of all plausibility structures with the same order n:*

a. *the total order (or the chain) has the minimal plausibilistic entropy $\hat{H} = 0$;*

b. *having established a fixed layer structure for D – the number of layers and the number of nodes in each layer – $(|L_0|, |L_1|, \ldots, |L_{l-1}|)$, the plausibilistic entropy is minimised when any two consecutive layers are minimally connected and maximised when any two consecutive layers are maximally connected:*

$$\frac{\sum_{k=0}^{l-2}\left(max\left(1, \frac{|L_{k+1}|}{|L_k|}\right) - 1\right)}{n} \leq \hat{H} \leq 1 - \frac{l}{n};$$

c. *the plausibilistic entropy of the fully connected structures strictly decreases when the number of layers increases;*

d. *the flat structure has the maximum entropy: $\hat{H}_n = 1 - \frac{3}{n}$;*

e. *any other structure (neither chain nor flat) has an entropy $\hat{H} \in (0, \hat{H}_n)$.*

Proof. **a.** The first observation to be made is that for any structure $\hat{H} \geq 0$. That is because $\forall v \in D$, $\perp \leq v \leq \top$ (according to Definition 3.1) so for any $v \neq \perp$ there will be an outgoing edge making $d_D^+(v) \geq 1$. The fact that the layer containing $\perp$ (bottom layer) is not taken in consideration when calculating $\hat{H}$ makes the numerator of the main fraction a sum of (not strictly) positive values. Since this amount is divided by a positive number the result can only be positive.

Now, considering a chain structure of order n, we have precisely n layers each of them with one element having exactly only one outgoing edge (except the bottom layer which does not count in this case) so $\hat{H}(chain_n) = \frac{\sum_{k=0}^{n-2}\left(\frac{1}{1}-1\right)}{n} = 0$.

b. The proof is trivial if we observe that any new edge in a fixed layered structure adds a new amount to a sum of positive values. The two formulas express the smallest and the largest number of edges for the plausibility structure to be well defined. For the upper limit we have

$$\frac{\sum_{k=0}^{l-2}\left(\frac{|L_k|\times|L_{k+1}|}{|L_k|} - 1\right)}{n} = \frac{\sum_{k=0}^{l-2}(|L_{k+1}| - 1)}{n} =$$

$$= \frac{(n-2)-(l-2)}{n} = 1 - \frac{l}{n}$$

c. Derived directly from **b.**

d. For $n = 2$ we have only one structure with two layers and $\hat{H}_n = 0$. For $n \geq 3$ the maximal entropy is reached when the number of layers is minimal (because of **c.**), that is $l = 3$, so $\hat{H}_n = 1 - \frac{3}{n}$.

e. Any non-chain structure will involve a strictly positive amount of choice that will render the numerator of the entropy strictly positive thus giving the structure a strictly positive plausibilistic entropy. On the other hand, any structure that is not a flat will not have the minimum number of layers possible and therefore its plausibilistic entropy will be strictly smaller than $\hat{H}_n$.

□

Proposition 3.3. *For any finite structure D,* $\hat{H}(D) \in [0,1)$.

Proof. For any finite plausibility set the minimum plausibilistic entropy equals 0 and is obtained when the structure is a chain (Proposition 3.3). On the other hand the plausibilistic entropy of a set D is maximised when the structure is flat and has the same number of elements. So, according to the same Proposition 3.3 no matter how large n may be its entropy will never reach 1.

Figure 4 shows the variation of the maximum plausibilistic entropy value of a structure depending on the number of the contained elements. □

We conclude with a small observation regarding the comparison of the top and the bottom plausibilistic entropies.

Proposition 3.4. $\hat{H} \neq \check{H}$.

Proof. For structure D_{31} in Figure 1, $\hat{H} = \frac{2}{7} \neq \frac{2+\frac{1}{3}}{7} = \check{H}$. □

Proposition 3.5. *Bottom plausibilistic entropy has the same properties as top plausibilistic entropy (Proposition 3.2) and in addition to that* $\hat{H} = \check{H}$ *for chains and flats.*

Proof. Obvious. Simply reverse the order relation so $\top$ and $\bot$ will be interchanged. □

3.3 Comparison with probabilistic entropy

Since a probability space is a special case of a plausibility space it would be interesting to find out what properties probabilistic and plausibilistic entropies have in common and how they differ.

The first difference that comes to mind is related to the properties that the corresponding support sets have. Specifically for the probability spaces the set of possible outcomes is a total order with elements supporting various operations (addition, multiplication, etc.). The elements of a plausibility space merely support a partial order. This makes it very difficult to model plausibilistic outcomes with probabilistic outcomes because there is no simple way of eliminating the rich properties of the latter while still maintaining the validity of the formulas. Therefore when comparing plausibility spaces with probability spaces we will limit ourselves to the plausibility structures that have fully connected layers as were used in Proposition 3.2 point c.

We will continue our analysis by defining the mapping between probability structures and plausibility structures that we will take into consideration. First the top and the bottom elements will be mapped the following way: $\top \leftrightarrow 1$ and $\bot \leftrightarrow 0$. Since it does not make sense to include these two elements in every probability structure the most simple plausibility structure that we will map is the one with three elements.

Probability structures are implicitly normalised in the spirit of Definition 3.2 and for distance we will use the "ordinary" Euclidean distance between individual probabilities. Therefore a layer in a probabilistic structure is the set of probabilities equally distanced from 1 (as $\top$). It is very easy to see that an alternative definition can be provided for layer in probability structures.

Definition 3.6 (layer)**.** *Given a probability structure P, by* layer *we understand the set of all the elements in P having the same probability.*

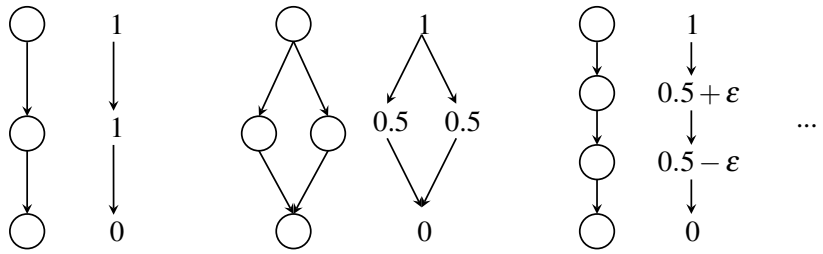


Figure 3: Mapping between plausibility and probability structures

The first obvious similarity between plausibilistic entropy and probabilistic entropy is related to the structure that maximises the entropy given an event space of a certain size n (note that if the plausibility structure will have n elements then the corresponding probability structure will have $n-2$ elements). According to Proposition 3.2 point c for plausibilistic entropy this configuration is achieved when all the elements (except for $\top$ and $\bot$) belong to the same layer. It is also very well known that the one layer structure is the one that also maximises the Shannon entropy [18]. Figure 4 allows for a visual comparison of the variations of the maximal values of the two entropies.

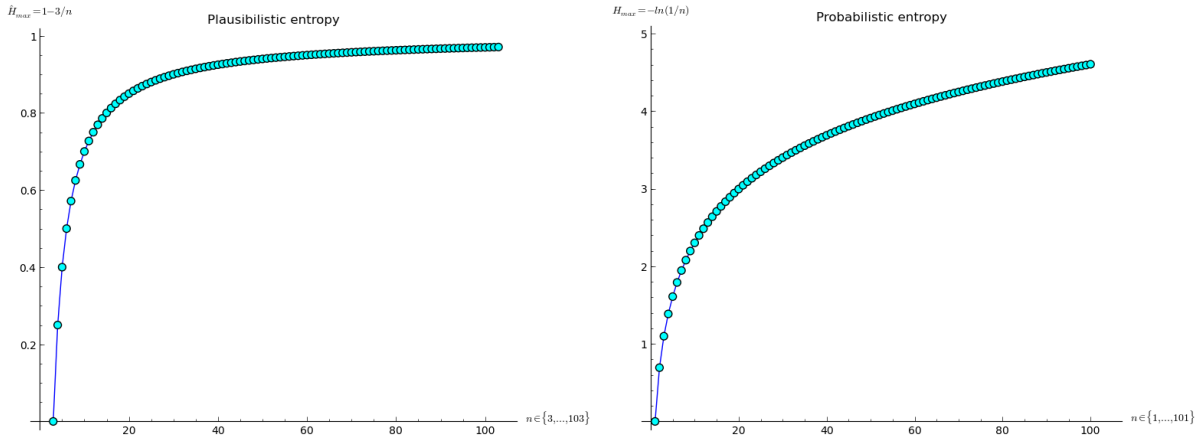


Figure 4: A comparison of the variation of the maximal entropy value for plausibilistic entropy and probabilistic entropy for a structure of size n.

Regarding the range of values that the two entropies can take we can note that plausibilistic entropy is convergent — $\lim_{n\to+\infty}\hat{H}_{max}=1$ (Proposition 3.3), while probabilistic entropy is not — $\lim_{n\to+\infty}H_{max}=+\infty$. For whatever meaningful size we can find a structure that minimises any of the two types of entropy to the same minimal value of 0. For plausibilistic entropy this happens when the structure is a chain and for the probabilistic entropy this happens when one probability is 1 and all others are 0 or non-existent.

The next question to be addressed is how does the number of layers affect the probabilistic entropy. Is this result similar to the one related to plausibilistic entropy stating that entropy strictly decreases when the number of layers increases? Given Definition 3.6 of a layer for a probabilistic structure it is very easy to understand what does it mean that the number of layers increases. On the other hand, given the continuous nature of the Shannon entropy, we have to make sure that the maximum value

possible of a k layered probability structure is greater than the maximum value possible of a $k+i$ layered $i>0$ probability structure of the same size n. In order to do so we will use some form of mathematical induction but first we will have to consider some intermediary results.

Proposition 3.6. *Given a set of $n>1$ values $P=\{p_1,p_2,\ldots,p_n\}$ with $p_i\in(0,1)$, and $\forall(i,j)\,p_i=p_j=p$, and $\sum_{i=1}^{n}p_i=k,k\in(0,1]$ then $\forall e>0$ and $e'=\frac{e}{n-1}$ and $P'=\{p_1+e,p_2-e',\ldots,p_n-e'\}$ the entropy of P' is smaller then the entropy of P.*

Proof. Basically we have to prove that

$$-n\cdot p\cdot log(p)>-(n-1)\cdot(p-e')\cdot log(p-e')-(p+e)\cdot log(p+e)$$

which by substitution and various operations translates to

$$\frac{k\cdot n\cdot log\left(-\frac{(e-k)\cdot n+k}{n^2-n}\right)+(e\cdot n+k)\cdot log\left(-\frac{e\cdot n^2-(e-k)\cdot n-k}{(e-k)\cdot n+k}\right)}{k\cdot n\cdot log\left(\frac{k}{n}\right)}<1.$$

We will now try to maximise the left hand side expression on e. By derivation on e we get

$$\frac{log\left(e+\frac{k}{n}\right)-log\left(-\frac{e}{n-1}+\frac{k}{n}\right)}{k\cdot log\left(\frac{k}{n}\right)}$$

having only one critical point for $e=0$. The second derivative for the same left hand side expression is

$$\frac{\frac{1}{e+\frac{k}{n}}-\frac{1}{(n-1)\cdot\left(\frac{e}{n-1}-\frac{k}{n}\right)}}{k\cdot log\left(\frac{k}{n}\right)}$$

which for $e=0$ evaluates to

$$\frac{\frac{n}{k}+\frac{n}{k\cdot(n-1)}}{k\cdot log\left(\frac{k}{n}\right)}$$

and since in our context this is a negative value it proves that a global maximum value was found for $e=0$. For $e=0$ the initial left hand side expression is 1 but since e is required to be strictly greater then 0 it follows that the strict inequality that we have to prove is true. □

Remark: This Proposition actually states that if we "move" an element from a layer in a probability structure, everything else remaining the same the entropy decreases.

Proposition 3.7. *Given a set of $n>1$ values $P=\{p_1,p_2,\ldots,p_n\}$ with $p_i\in(0,1)$, and $\forall(i,j)\,p_i=p_j=p$, and $\sum_{i=1}^{n}p_i=k,k\in(0,1]$ then for any transformation that creates a two layered structure by removing $m<n$ elements from P and placing them at distance e, there is another transformation that takes only one element from P and also creates a two layered structure at distance e' the entropies of the two structures resulted from transformation being equal. This happens taking into consideration that if the original e is small enough.*

Proof. We have to prove that an e' exists such that $H_m=H_1$ where

$$H_m=-m\cdot\left(\frac{k}{n}+e\right)\cdot log\left(\frac{k}{n}+e\right)-(n-m)\cdot\left(\frac{k}{n}-\frac{m\cdot e}{n-m}\right)\cdot log\left(\frac{k}{n}-\frac{m\cdot e}{n-m}\right),$$

and

$$H_1=-\left(\frac{k}{n}+e'\right)\cdot log\left(\frac{k}{n}+e'\right)-(n-1)\cdot\left(\frac{k}{n}-\frac{e'}{n-1}\right)\cdot log\left(\frac{k}{n}-\frac{e'}{n-1}\right).$$

We notice that

$$\lim_{e\to 0} H_m(e) = -k \cdot log\left(\frac{k}{n}\right) = \lim_{e'\to 0} H_1(e')$$

and since the value in the middle is well defined ($k > 0$ and $n > 0$) we have the proof that when $e \to 0$, e' must also approach 0 so it exists. Also e' cannot be 0 because otherwise e should also be 0 and that will contradict the hypothesis. □

Remark: This proposition proves that splitting a layer in two in a probability distribution by taking any number of elements apart to create the two layered structure is equivalent entropy-variation-wise with operating upon a single element. This means that when operating strictly on a single layer to increase the number of layers in a structure the entropy always decreases.

Proposition 3.8. *Given a discrete probability distribution $P = \{p_1, p_2, \ldots, p_n\}$ with at least one layer L, $|L| > 1$, and all the elements in L having the same value $p \neq 0$, for every distance $d_1 > 0$, d_1 sufficiently small, that is used to create a new layer at distance d_1 from L by taking an element from L and modifying all the other probabilities accordingly, there is a value d_2 that can specify the distance between the moved element and the remaining elements in L such the entropy will be preserved (same to the one obtained by the previous transformation) while not affecting the values of any element that is not in L.*

Proof. To save space only a sketch of the demonstration will be presented here.

Let $L = \{\overline{p_1}, \overline{p_2}, \ldots, \overline{p_l}\}$ be the layer to be split and $\overline{p}$ the element that we are going to operate upon.

For the first case we have a set of values $E^* = \{\mathbf{e}^*\} \cup \{\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_{l-1}^*\} \cup \{e_1^*, e_2^*, \ldots, e_{n-l}^*\}$ that will have to be added to the elements in P to obtain the new distribution P^*. We will thus have $d_1 = |p + \mathbf{e}^* - p - \mathbf{e}_1^*| = |\mathbf{e}^* - \mathbf{e}_1^*|$.

In the second case we will have $E' = \{\mathbf{e}'\} \cup \{\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_{l-1}'\}$ that will have to be added to the elements in L to obtain the new distribution P'. We will thus have $d_2 = |p + \mathbf{e}' - p - \mathbf{e}_1'| = |\mathbf{e}' - \mathbf{e}_1'|$.

What we have to prove is that $(\forall d_1 \to 0)(\exists d_2)(H(P^*) = H(P'))$.

Because the sum of probabilities must always be 1 we can prove that

$$\lim_{d_1\to 0} H(P^*) = H(P) = \lim_{d_2\to 0} H(P')$$

and we can conclude that the Proposition 3.8 is true. □

Now we can prove the following similarity between plausibilistic entropy and probabilistic entropy.

Proposition 3.9. *For finite probability spaces increasing the number of layers on a fixed size distribution leads to a decrease in the maximal entropy that can be obtained by that structure.*

Proof. We will prove this by induction on the number of layers.

As a first step let us assume that we have a probability distribution with only one layer. From literature this is known to be the structure with the highest probabilistic entropy. Then obviously any other structure has a smaller entropy and so does a two layered structure too. This proves the base step.

Now we have to prove that given a layered discrete probability distribution that aims to maximise the entropy, by increasing the number of layers we get a new structure whose maximised entropy is smaller then that of the previously described structure.

To better explain what we understand by layered structure that maximises the entropy let us consider that we have a discrete probability distribution with n elements $P = \{p_1, p_2, \ldots, p_n\}$ divided into k layers of sizes $l_1, l_2, \ldots, l_k$. It should also be noted that $\sum_{i=1}^{n} p_i = 1$ and $\sum_{i=1}^{k} l_i = n$. Since $p_i = p_j$ if p_i and p_j belong to the same layer we can introduce the aforementioned probability distribution by considering a reference value $p = max(P)$ and a set of values $E = \{e_1, e_2, \ldots, e_{k-1}\}$ representing the distances between

consecutive layers. For convenience we can introduce an additional value $e_0 = 0$ and say that if for some $p' \in P$ belonging to layer i, $p' = p + e_0 + e_1 + \cdots + e_{i-1}$. It is now easy to see that the entropy of such a structure is maximised when all the values an E approach 0 simultaneously. Still, since none of the values in E can actually be 0 (this will lead to a decrease in the number of layers) the maximal value of the layered structure entropy will not reach the absolute maximum value of the entropy obtained only by the single layer structure.

Given a k layered probability structure let us take one element from a layer and create a $k+1$ layered structure. Considering the restrictions imposed by the context of our demonstration first of all we can only take that element from a layer that already has at least two elements. Because the sum of of the individual probabilities must always be 1 the fact that we "move" an element from a layer in order to create a new layer will implicitly "affect" the values of the elements from at least another layer. Finally, letting e be the value that we add to the element that we choose to form a new layer, we can conveniently choose $e \neq 0$ in such a way that the new $k+1$ layered structure is well defined.

Let us consider the case where the only "affected" layer is the one that we take the element from. In this case, sice all other probabilities remain the same, according to Proposition 3.6 the entropy decreases. If we take more than one element from a layer to form a new layer then the entropy will also decrease because according to Proposition 3.7 creating a new layer from more then one element is equivalent to creating the new layer from only one element and we have just established this to be an entropy lowering operation. The fact that the child layer must be at a very small distance to the parent layer not only respects the conditions of our lemma but also maximises the entropy of the newly formed structure.

Finally, if the movement of an element affects more than one layer then, according to Proposition 3.8, this is an operation equivalent to affecting only one layer entropy-variation-wise and we have already established that this has a lowering effect on the entropy. Choosing to move any number of elements from a layer in order to create a new one while affecting the entire structure in the process can also be proven to be reducible to operating on only one element.

Thus we can conclude that in the specified conditions the maximal entropy that can be obtained by a (k)-layered structure is higher then the maximal entropy that can be obtained by a $(k+1)$-layered structure. □

Remark: Proposition 3.9 does not say that increasing the number of layers in a probability structure will always decrease the entropy, for instance $H(\{0.325;0.33;0.345\}) \sim 1.98 > 0.112 \sim H(\{0.01;0.01;0.98\})$, it says that when we try to obtain the maximal value of the entropy for a probability distribution we should do our best to minimise the number of layers. This is important because when it is difficult for us to differentiate between probabilities (hard to grasp infinitesimal differences) then using the plausibilistic approach can help us make a decision. On the other hand, if we can clearly differentiate between probabilities then using the plausibilistic approach loses its edge. Still, the distance between two probabilistic layers could be somehow intuitively related to the degree of connectivity between plausibilistic adjacent layers and the similarity between the two approaches could be pushed further. A greater distance between probabilistic layers could mean a lower connectivity between plausibilistic layers an in both cases that is translated to a decrease in entropy.

The main result of this comparison is the proof of a certain degree of consistency between the concept of plausibilistic entropy and that of probabilistic entropy. While we cannot talk about plausibilistic entropy as being a generalisation for Shannon entropy, the current results do not exclude the possibility of using it as an alternative, especially when the actual numbers are difficult or even impossible to obtain. In one of the following sections, 4.3, we will provide a potential application for this concept.

	Plausibilistic entropy	**Probabilistic entropy**
Top element	$\top$	1
Bottom element	$\bot$	0
Properties	$<$	$\mathbb{R}$
Layer definition	a set of all the elements equally distanced from $\top$	a set of all the elements having the same probability
Formula	$\hat{H} = \frac{\sum_{k=0}^{l-2} \left(\frac{\sum_{v \in L_k} d_D^+(v)}{\lvert L_k \rvert} - 1 \right)}{n}$	$H = -K \sum_{i=1}^{n} p_i \log(p_i)$
Domain (discussed here)	finite plausibility structures	discrete probability distributions
Codomain	$[0,1)$	$[0,+\infty)$
Maximum entropy structure	flat structure	uniform probability distribution
Maximum entropy value	$1-\frac{3}{n}$	$-log\left(\frac{1}{n}\right)$
Minimum entropy structure	chain structure	$\{1,0,\ldots,0\}$
Minimum entropy value	0	0
Structural correlations (fixed size structures)	for fully connected structures entropy strictly decreases if the number of layers increases	the maximal value of the entropy that a structure can obtain decreases when the number of layers increases
Applicability	when obtaining the actual probability values is impossible or too expensive	whenever we have enough information to make educated guesses regarding the probabilities

Table 1: A side-by-side comparison between plausibilistic entropy and probabilistic entropy.

3.4 Conditioning plausibilities

We would like to close this discussion about plausibilistic entropy with some concepts related to conditioning plausibilities that will both extend the generalisation and help us in providing quantitative definitions for information security concepts. We are talking about *conditional plausibility measures* (CPMs) and *conditional plausibility spaces* (CPSs).

Definition 3.7 (conditional plausibility space [11])**.** *A* conditional plausibility space *(CPS) is a tuple* $(\Omega, \mathscr{F}, \mathscr{F}', \nu)$ *with:*

Ω *is a set of possible outcomes;*

$\mathscr{F} \times \mathscr{F}'$ *is a Popper algebra over* Ω*:*

- $\mathscr{F}$ *is an algebra over* Ω*;*
- $\mathscr{F}' \neq \emptyset$ *with* $\mathscr{F}' \subseteq \mathscr{F}$*;*
- $\mathscr{F}'$ *is closed under supersets in* $\mathscr{F}$*: if* $V \in \mathscr{F}'$*,* $V \subseteq V'$*, and* $V' \in \mathscr{F}$ *then* $V' \in \mathscr{F}'$*;*
- *if* $U \in \mathscr{F}$*,* $V \in \mathscr{F}'$*, and* $\nu(U|V) \neq \bot$ *then* $U \cap V \in \mathscr{F}'$*.*

ν *is a* conditional plausibility measure *(CPM) mapping* $\mathscr{F} \times \mathscr{F}'$ *to* plausibility structure D *with the following properties:*

1. $\nu(\emptyset|V) = \bot$*;*
2. $\nu(\Omega|V) = \top$*;*
3. *if* $U_1 \subseteq U_2$ *then* $\nu(U_1|V) \leq \nu(U_2|V)$*;*
4. $\nu(U|V) = \nu(U \cap V|V)$*.*

Remark: Since the elements of the plausibility spaces are not required to elicit the reach properties of the real numbers, the concept of conditional plausibility cannot be derived by a formula, as for conditional probabilities, it has to be introduced by using low level concepts from set theory.

4 Anonymity: a short formal analysis

4.1 Qualitative definition

The intuition behind the concept of secrecy is, as Halpern and O'Neil formulated in [5]: "one agent maintains secrecy with respect to another if the second agent cannot rule out any possibilities for the behaviour or state of the first agent". They also gave formal definitions to characterise various degrees of secrecy. In this section we will focus on some specific forms of secrecy – anonymity (hiding the Actor). We agree with the affirmation in [6] stating that "the basic intuition behind anonymity is that actions should be divorced from the agents who perform them for some set of observers".

Definition 4.1 (minimal anonymity [17])**.** *An action a, performed by an agent A, is* minimally anonymous *in* $\mathscr{I}$ *w.r.t. an agent I if*

$$\mathscr{I} \models \theta(A,a) \Rightarrow P_I[\neg\theta(A,a)].$$

Definition 4.2 (group anonymity [17])**.** *Action a, performed by agent A, is* anonymous up to anonymity set $G \subseteq Ag \setminus \{I\}$ *in* $\mathscr{I}$ *w.r.t. agent I if*

$$\mathscr{I} \models \theta(A,a) \Rightarrow \forall_{(A' \in G)} P_I[\theta(A',a)].$$

Remark: From a more intuitive point of view, minimal anonymity simply states that a third party (e.g. the observer) cannot tell with absolute certainty that some action was or was not performed by a certain agent. No qualitative assessment is required. On the other hand, in the case of the group anonymity, it is required that the observer should suspect that any of agents in a group may have performed a certain action. If the size of the group is 1 then the observer will definitely know who the action performer was and we cannot accept this to be an authentic anonymity. The relation between the two definitions is rather complex. While minimal anonymity focuses on the core of what anonymity is (not knowing who did something), group anonymity suggests a way of achieving/realising this idea of anonymity. Group anonymity also hints to the idea of a spectrum of degrees of anonymity ranging from 0 (when the anonymity set has only one element thus allowing the observer to know exactly the doer) to the maximum degree possible – intuitively the more suspects the greater the amount of anonymity. In fact, for some

special cases (when there are at least two agents in the anonymity set), it can be formally proven that group anonymity implies minimal anonymity [6]. Analysing how the number of agents in an anonymity group and the information an observer has about them affect the degree of anonymity will be the focus of the next section.

4.2 Quantitative definition

4.2.1 Probabilistic approach

In [6] Halpern and O'Neil tackle the problem of measuring anonymity by using probabilities. Their approach starts with a probability distribution over the set of runs and, by using the Halpern-Tuttle construction [19], derive a point based probability distribution capable of expressing the fact that a certain agent A assigns a certain probability $Pr_A(\varphi)$ to a formula φ. In this context the following two probabilistic definitions for anonymity are proposed:

Definition 4.3 (probabilistic $\alpha-$anonymity [6])**.** *Given a value $\alpha \in (0,1]$, action a, performed by agent A, is* $\alpha-$anonymous *w.r.t. agent I if*

$$\mathscr{I} \models Pr_I[\theta(A,a)] < \alpha.$$

Definition 4.4 (strong probabilistic group anonymity [6])**.** *Action a, performed by agent A, is* strongly probabilistically anonymous up to anonymity set $G \subseteq Ag \setminus \{I\}$ *in $\mathscr{I}$ w.r.t. agent I if for each $A' \in G$*

$$\mathscr{I} \models Pr_I[\theta(A,a)] = Pr_I[\theta(A',a)].$$

These are obvious quantitative counterparts of the anonymity definitions introduced previously. While very powerful, they fail to provide an overall view of the situation an observer is in relative to a decision making perspective. As the authors themselves mentioned, the former case does not provide a relevant measure for anonymity for any α and any probability distribution — however small we choose α to be (e.g. 0.01) we can always find a probability distribution that contains one element with a probability that is very close to α and a very large number of other elements (e.g. $10,000$) equally dividing the rest of the probability. This would make it very easy for the observer to choose a possible actor. The latter definition is way too strong because it does not allow for small, practically irrelevant, variations in the probabilities to exist.

To overcome this limitation we could try an approach that uses a unique measure capable of characterising a probability distribution. An excellent such measure is the entropy. In [8] Deng et al. employ relative entropy to define anonymity. While their formalism differs from the one presented here we believe that there are ways to introduce some form of entropy [20, 15] based anonymity definition in our framework.

The classic entropy based approach still suffer from a significant limitation: the actual numbers can be difficult to obtain; and therefore we will move on to a more flexible approach.

4.2.2 Plausibilistic approach

We start with a plausibility distribution over the set of runs of a MAS and, by using the Halpern-Tuttle construction [19, 7], we obtain a point based plausibility distribution. The details for this are presented in the ARES paper [1].

We consider $Pl_A(\varphi)$ to be the plausibility that agent A attaches to a formula φ at point (r,m). Now we could define some form of α plausibilistic anonymity requiring that $\mathscr{I} \models \theta(A,a) \Rightarrow Pl_I(\theta(A,a)) \leq \alpha$, $\alpha \in D$. However, since D is merely a partially ordered set, this requirement could severely restrict the number of the real situations this theory could be taken applied to.

A plausibilistic variant of the strong probabilistic anonymity up to a certain anonymity set G would require equal plausibilities for all contained agents: $\mathscr{I} \models \theta(A,a) \Rightarrow \forall_{(A',A''\in G)} Pl_I(\theta(A',a)) = Pl_I(\theta(A'',a))$. This variant however is a simple theoretical construction that brings nothing new relative to the probabilistic approach.

By using the plausibilistic entropy we could generalise the idea of strong plausibilistic anonymity:

Definition 4.5 (plausibilistic α group anonymity [1])**.** *Action a, performed by agent A, is* plausibilistically α-anonymous up to anonymity set $G \subseteq Ag \setminus \{I\}$ *in the plausibilistic interpreted system* $\mathscr{I}$ *w.r.t. agent I if*

$$\mathscr{I} \models \theta(A,a) \Rightarrow \hat{H}\left(\{Pl_I[\theta(A',a)] : A' \in G\}\right) \geq \alpha.$$

Remark: This form of defining anonymity takes into consideration all the information an observer has about the agents in an anonymity set. This form of anonymity is well defined regardless of how I structures the information it possesses regarding the agents in the anonymity set. Because α can take any rational value in $[0,1)$ any two degrees of anonymity can be compared, meaning that given two anonymity providing protocols they can be compared based on the level of plausibilistic anonymity they provide. The anonymity roughly grows with the number of the agents in the anonymity set. Supporting the intuition is the fact that while 0 plausibilistic group anonymity can be easily obtained absolute anonymity ($\alpha = 1$) cannot be provided by a finite anonymity set [1].

4.3 Application

In [14] Goel et al. introduced a DC-networks based anonymity protocol they called Herbivore. The scalability of the Herbivore protocol is ensured by partitioning the network into dynamically allocated anonymising cliques. Herbivore guarantees that each clique will have at least k nodes, where k is a predetermined constant assumed to describe the degree of anonymity provided by the system. Relevant to our discussion is the algorithm employed to allocate the cliques: new cliques are created automatically when existing cliques grow too large to communicate efficiently and when the number of nodes in a clique falls below k, the nodes in that clique are redistributed throughout the network. Unrelated to any anonymity quantifying theory, in their reference implementation, the authors chose 64 as a value for k. According to the authors, adversaries able to monitor all network traffic cannot deduce the identity of a sender or receiver beyond an anonymising clique. This means that an observer could identify the clique that originated a message and/or the destination clique but have no idea related to the identity of the actual agents that communicated.

This relates to the theory that we present here in the following manner: we consider the participants in a clique to form an anonymity group G as used in Definitions 4.2, 4.4, and 4.5. It is easy to see that, at a clique lever for $k \geq 2$, the Herbivore protocol provides the property of group anonymity as it is qualitatively defined. The authors actually use k as a measure for the degree of anonymity the protocol provides and this is relevant as long as the conditions of the strong probabilistic group anonymity apply. In real life however [21], an observer can analyse the message itself in order to extract information that could identify the communicating parties and, it is our belief, that this should be factored in when designing an anonymity protocol. Particularly the Herbivore protocol could be extended to use a well defined anonymity measure instead of just k when having to make the decision of reshuffling the cliques.

By using our definition of plausibilistic anonymity the original value of k could be transformed into a reference value of plausibilistic entropy by using the formula introduced by Proposition 3.2 point d. (e.g. for $k = 64$, $\hat{H}_k = 0.953125$) and then invoke the clique reshuffling procedure when the anonymity gets below this value. Aware of the difficulties such an approach would entail, what we try to say here is that anonymity protocols in general could take advantage of an elaborated definition for quantitative anonymity and, due to its flexibility, plausibilistic group anonymity could be a valuable candidate.

5 Conclusion

The main focus of this paper is the in-depth analysis of the *plausibilistic entropy*. Plausibilistic entropy aims to be for plausibility spaces what Shannon entropy is for probability spaces: a simple measure for the degree of choice or uncertainty existing in a system. This new concept was introduced in order to allow for a basic quantitative characterisation of information security properties, like anonymity, in the situations where the precise numeric values of the probabilities are not available.

We begin by summarising the theoretical framework that we employ for reasoning about security properties. Basically, what we use is an epistemic logic defined in a MAS setting. The dynamic behind is that various agents, communicating with one another, want their activity to be hidden from a certain third party. The formal framework that is introduced helps us to rigorously express the statements characterising the security related properties of the interactions.

Afterwards, in order to quantitatively express the degree to which the security properties are satisfied, we analyse the plausibilistic entropy. This concept allows for a very high level of generality to be used for expressing uncertainty, because the constructs that it is based upon are very basic: plausibility structures. Plausibility structures only require a partial order relation to be assumed and this is enough for us to device a way that allows us to quantify the amount of choice in the system. Among the most practical properties of the plausibilistic entropy is the one allowing the comparison of any two plausibility structures, no matter how complex, by managing to map them into the $[0, 1)$ interval. The main contribution of this paper consists of the side-by-side analysis between the plausibilistic entropy and the classical probabilistic entropy. Based on this analysis, our conclusion is that, while we cannot assume (for now) a generalisation relation between the two of them, there are compatibilities between the two notions to such a degree that the concept of entropy could be extended to structures that are more general then probability spaces. Such an investigation could be the topic of a further research.

Last, we attempt the compilation of a short list of the anonymity definitions compatible with our logical framework, both qualitative and quantitative, and demonstrate how the plausibilistic definition of anonymity can be used to enhance the understanding of existing anonymity protocols like Herbivore.

Acknowledgement

The author would like to thank Ferucio Laurenţiu Ţiplea Professor, Ph.D. for reviewing the current version of the material and for providing the encouragements that led to its development.

References

[1] I. Goriac, "Measuring anonymity with plausibilistic entropy," in *Proc. of the 8th International Conference on Availability, Reliability, and Security (ARES'13), Regensburg, Germany*. IEEE, September 2013, pp. 151–160.

[2] ——, "An epistemic logic based framework for reasoning about information hiding," in *Proc. of the 6th International Conference on Availability, Reliability, and Security (ARES'11), Vienna, Austria*. IEEE, August 2011, pp. 286–293.

[3] ——, "Compiling an epistemic logic for multiagent systems," *International Journal of Intelligent Systems*, vol. 28, no. 7, pp. 648–668, July 2013.

[4] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi, *Reasoning about Knowledge*. The MIT Press, 2004.

[5] J. Y. Halpern and K. R. O'Neill, "Secrecy in multiagent systems," *ACM Transactions on Information and System Security*, vol. 12, no. 1, pp. 1–47, October 2008.

[6] ——, "Anonymity and information hiding in multiagent systems," *Journal of Computer Security*, vol. 13, no. 3, pp. 483–512, May 2005.

[7] K. R. O'Neill, "Secrecy and anonymity in interactive systems," Ph.D. dissertation, Cornell University, August 2006.

[8] Y. Deng, J. Pang, and P. Wu, "Measuring anonymity with relative entropy," in *Proc. of the 4th International Workshop on Formal Aspects in Security and Trust (FAST'06), Hamilton, Ontario, Canada, LNCS*, vol. 4691. Springer Berlin Heidelberg, August 2007, pp. 65–79.

[9] J. Y. Halpern, *Reasoning about Uncertainty*. The MIT Press, 2005.

[10] ——, "Conditional plausibility measures and bayesian networks," *Journal of Artificial Intelligence Research*, vol. 14, no. 1, pp. 359–389, January 2001.

[11] ——, "Plausibility measures: A general approach for representing uncertainty," in *Proc. of the 17th International Joint Conference on Artificial Intelligence (IJCAI'01), Seattle, Washington, USA*, vol. 2. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, August 2001, pp. 1474–1483.

[12] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[13] W. A. Stein and others / The Sage Development Team, "Sage mathematics software (version 5.12)," 2013, `http://www.sagemath.org`. [Online]. Available: http://www.sagemath.org

[14] S. Goel, M. Robson, M. Polte, and E. G. Sirer, "Herbivore: A scalable and efficient protocol for anonymous communication," Cornell University, Tech. Rep. 1890, February 2003.

[15] C. Cachin, "Entropy measures and unconditional security in cryptography," Ph.D. dissertation, Swiss Federal Institute of Technology Zürich, 1997.

[16] R. Diestel, *Graph Theory*. Springer, 2010.

[17] Y. Tsukada, K. Mano, H. Sakurada, and Y. Kawabe, "Anonymity, privacy, onymity, and identity: A modal logic approach," *Transactions on Data Privacy*, vol. 3, no. 3, pp. 177–198, December 2010.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[19] J. Y. Halpern and M. R. Tuttle, "Knowledge, probability, and adversaries," *Journal of the ACM*, vol. 40, no. 4, pp. 917–962, September 1993.

[20] A. Renyi, "On measures of entropy and information," in *Proc. of the 4th Berkeley Symposium on Mathematics Statistics and Probability, Berkeley, Calif., USA*, vol. 1. University of California Press, June–July 1961, pp. 547–561.

[21] M. Backes, G. Doychev, M. Dürmuth, and B. Köpf, "Speaker recognition in encrypted voice streams," in *Proc. of the 15th European Symposium on Research in Computer Security (ESORICS'10), Athens, Greece, LNCS*, vol. 6345. Springer Berlin Heidelberg, September 2010, pp. 508–523.

[22] T. Jech, *Set Theory*. Springer, 2006.

[23] F. L. tiu Ţiplea, L. V. amanu, and C. V. arlan, "Complexity of anonymity for security protocols," in *Proc. of the 15th European Symposium on Research in Computer Security (ESORICS'10), Athens, Greece, LNCS*, vol. 6345. Springer Berlin Heidelberg, September 2010, pp. 558–572.

[24] ——, "Reasoning about minimal anonymity in security protocols," *Future Generation Computer Systems*, vol. 29, no. 3, pp. 828–842, March 2013.

[25] W. P. Webber and L. C. Plant, *Introductory Mathematical Analysis*. HardPress Publishing, 2013.

Author Biography

Iulian Goriac is a spare time researcher interested in computer science, economics, and psychology. He finances his studies by working as a full time software developer. You may learn more about him by browsing his site (`https://sites.google.com/site/mindtripshome`) or by writing him an email (iulian.goriac@gmail.com).