An information-theoretic model for steganography (2003 IH)

▼ 模型假设

▼ 故事背景

Alice 和 Bob 被关进监狱，分别被锁在相距甚远的单独牢房内，他们希望制定一个**越狱计划**。

他们被允许通过**可信使者**发送**经过认证的消息**进行交流，但前提是这些消息**不能涉及越狱计划**。

然而，使者其实是**典狱长 Eve（对手）的特工**，她会**窃听**他们所有的通信内容。

如果 Eve 发现任何**密谋的迹象**，她将立刻采取行动，把 Alice 和 Bob 关进**最高安全级别的牢房**，从那里**从未有人成功逃脱**。

▼ 囚徒的对策

Alice 和 Bob 非常清楚这一点，因此在入狱之前，他们**共享了一个秘密代码词**。现在，他们计划利用这个代码词，在**看似无害的信息**中**隐藏真正的含义**，以此秘密交流。

如果 Alice 和 Bob **成功交换信息**，让他们能够协调行动，而 Eve **没有察觉到任何异常**，那么他们的计划就成功了。

▼ Eve 的挑战

- Eve 知道囚犯之间通常会进行哪些**正常通信**，这种通信被称为正常文本（covertext）。
- **她同样知道囚犯可能会使用某些技巧，在表面上无害的消息中**隐藏秘密信息，这样的消息被称为“隐写文本”（stegotext）。
- **按照**信息论的方法，我们用**概率模型**来描述 Eve 的知识，并将她检测隐藏消息的任务视为**假设检验（hypothesis testing）问题**。

▼ 隐写系统的安全性定义

我们用 **相对熵（或称判别信息）** 来衡量 **正常文本分布** 和 **隐写文本分布** 之间的区别。

当**相对熵为零**时，我们称该隐写系统为**完美隐写系统（perfect stegosystem）**。

▼ 核心重点

- 攻击者知道正常文本分布
- 通过相对熵来判断是否安全，当**相对熵为零**时，我们称该隐写系统为**完美隐写系统**
- 这是个**对称隐写系统**, 隐写双方共享了一个秘密代码词

▼ 数学定义

▼ 1. 相对熵（KL 散度）的回顾

- 隐写发布者需要计算相对熵, 来猜测监管者的分辨能力
- 相对熵（也称 Kullback-Leibler 散度, KL divergence）是衡量两个概率分布 P 和 Q 之间的不同程度的一种方法。它表示如果我们使用已知分布 Q 来近似分布 P，我们会损失多少信息。
- 相对熵（KL 散度）衡量两个概率分布之间的差异，定义为：

$$D(P\|Q) = \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)}$$

其中：

- P(x) 是实际分布（真实分布）。
- Q(x)是对比分布（通常是隐写分布）。
- 如果 P(x) 和 Q(x) 完全相同，则相对熵 D(P∥Q)=0
- 如果 P(x) 和 Q(x) 差异越大，则 D(P∥Q) 也越大。

▼ 2. 条件相对熵（Conditional Relative Entropy）定义

$$D(P_{Q_0|V}\|P_{Q_1|V}) = \sum_{v\in\mathcal{V}} P_V(v)\sum_{q\in\mathcal{Q}} P_{Q_0|V=v}(q)\log\frac{P_{Q_0|V=v}(q)}{P_{Q_1|V=v}(q)}$$

- **核心思想**：条件相对熵衡量的是在**已知某个额外条件（或变量）V 的情况下**，两个（条件）概率分布之间的差异。
- **对比普通相对熵**：

- 普通的相对熵直接衡量两个概率分布$P(X)$和$Q(X)$ 的不同。
- 条件相对熵则是在引入一个条件变量 V 后，衡量 $P(X|V)$和 $Q(X|V)$ 这两个条件概率分布之间的不同。

- **内部求和**

$$\sum_{q \in \mathcal{Q}} P_{Q_0|V=v}(q) \log \frac{P_{Q_0|V=v}(q)}{P_{Q_1|V=v}(q)}$$

  - 这部分计算的是**在给定条件 V=v 的情况下**，两个条件概率分布 $P_{Q_0|V=v}$ 和 $P_{Q_1|V=v}$ 之间的相对熵。

- **外部求和**

$$\sum_{v \in \mathcal{V}} P_V(v)[\dots]$$

  - $P_V(v)$ 是条件变量 V 取值为 v 的概率。
  - 整个外部求和表示对所有可能的条件 v，将该条件下Q_0 和 Q_1 的相对熵，按照条件 v 发生的概率$P_V(v)$进行加权平均。
  - 可以理解为：我们在知道额外信息 V 后，重新计算两个分布之间的差异，然后对所有 V 取值求加权平均。

▼ 例子

我们可以这样更深入地理解条件相对熵：

想象一下，我们试图判断一个“载体”（比如一封邮件、一张图片、一段音频，或者更广义的数字内容）是否隐藏了秘密信息。这个“载体”本身具有极其丰富的、**多维度的特征**。这时，我们所说的额外信息V就不再仅仅是一个单一的、孤立的属性（比如“邮件主题”），而是可以代表这个载体在多个不同维度上的具体表现或状态的集合

例如，对于一个图像载体，V 可以是：

- **视觉**: 图像的整体亮度分布、对比度、颜色直方图等视觉层面的一阶统计特征。
- **频域**: 图像在DCT变换域、小波变换域等特定频域的系数分布特征。
- **空域邻近性**: 像素与其邻近像素之间的相关性、局部纹理复杂度等。

- **模型特定**: 由某个预训练的深度学习模型（如CNN的中间层激活值）提取出来的高维抽象特征。
- **内容语义**: 图像所描绘的场景类型（室内、室外、人像、风景等），这本身也可能影响其底层统计特性。

对于文本载体，V 可以是：

- **词汇**: 词频分布、特定关键词的出现模式。
- **句法**: 句子长度分布、依存关系模式、词性序列模式。
- **语义**: 文本的主题类别、情感倾向。
- **生成路径**: (在生成式隐写中) 生成该文本时，模型在某些关键决策点的选择倾向或内部状态。

条件相对熵的作用在于

- 我们先“聚焦”到这些多维度特征中的**某一个或某一组特定特征组合**上（即我们“知道”了V的具体取值，比如“这张图片的亮度分布是A，并且其DCT系数的某个统计量是B”）。
- 然后，**在这个特定的“视角”或“条件下”**，我们再来比较“原始无秘密信息的载体”（其内容分布为 Q_0）和“可能嵌入了秘密信息的载体”（其内容分布为 Q_1）在**其他可观察特征或整体表现**上的分布差异。
- 最后，我们会对所有我们关注的V的可能取值
- 即所有我们考虑的多维度特征组合的可能性进行加权平均，根据V发生的概率，得到一个综合的、在该条件V下的平均差异度量。

这就像是我们只选择了一个非常特定的 V（比如，仅仅是图片整体像素值的一阶统计量，或者邮件的“主题”这一个维度），然后发现条件相对熵$D(P_{Q_0|V} \| P_{Q_1|V})$在这个**单一维度V**下趋近于0。我们可能会据此声称系统在这个维度上是“安全”的。

> “但这是否就意味着系统在所有维度上都是安全的呢？
>
> VAE-Stega等研究给我们敲响了警钟：当一个维度的指标被极度优化（比如失真被控制到极小）时，很可能导致其他维度的安全指标不尽如人意，甚至出现劣化。”

- 如果我们只关注了维度 $V_{\text{甲}}$（例如，努力让修改后的图片在“视觉失真”这个维度 $V_{\text{甲}}$ 上，其条件 KL 散度相对于原始图片极小，即

$D(P_{Q_0|V_{甲}} \| P_{Q_1|V_{甲}}) \approx 0$，那么秘密信息可能会以一种在 $V_{甲}$ 维度上不明显

- 但在另一个维度$V_{乙}$，例如，“DCT系数的特定高频分量的统计特性”，或者“邻近像素差分直方图”）上非常明显的方式被编码进去。
- 此时，如果我们计算基于维度$V_{乙}$的条件相对熵 $D(P_{Q_0|V_{乙}} \| P_{Q_1|V_{乙}})$，它可能会非常大，表明在这个 $V_{乙}$ 维度上，隐写信息极易被检测。
- 这就是VAE-Stega等研究指出的现象：**为了在一个维度（如失真）上做到极致，可能会在另一个维度（如某种统计可检测性）上暴露出更大的破绽。**

因此，这个直观解释可以进一步深化为：

条件相对熵允许我们**有选择性地、针对性地**考察在某些**特定维度或特征组合 (V)** 的条件下，隐写操作对载体其他统计特性 (Q) 或整体分布的影响。

一个真正鲁棒和安全的隐写系统，其目标不应仅仅是在少数几个、甚至单一预设的维度 (V) 上获得较低的条件相对熵。

更理想的情况是，无论攻击者从哪个合理的特征维度V（或维度组合）去观察，嵌入秘密信息后的载体分布$P_{Q_1|V}$ 都应该与原始载体分布 $P_{Q_0|V}$足够接近，即相应的条件相对熵都应该足够小。

- VAE-Stega的警示告诉我们，隐写算法的设计和评估需要一种**多维度的、整体的视角**。
- 过度优化一个维度上的条件相对熵（使其趋近于0），如果导致了在其他未被充分考虑的维度被忽略上的条件相对熵显著增大，那么这种“单点最优”可能是一种“假安全”或“脆弱的安全”。
- 真正的安全需要在多个重要且可能被攻击者利用的特征维度之间取得平衡和鲁棒性。
- 条件相对熵这个工具，如果被恰当地应用于分析多个不同维度的V，就能帮助我们更好地理解隐写系统在多维特征空间中的安全性，并揭示这种维度间的制衡与取舍关系。

▼ 隐写系统交互模式定义

- 在隐写系统中，存在两个用户，Alice 和 Bob。

- Alice 希望通过**公共信道**向 Bob 发送**看似无害但实际包含隐藏信息的消息**，同时使得第三方攻击者 Eve **无法察觉隐藏信息的存在**。
- Eve 被认为具有**完美的只读访问权限**，可以完全监听公共信道上的所有通信内容。

▼ Alice 的两种操作模式

Alice 的操作模式有两种：

1. Alice 处于非活跃状态（无隐写信息）

- Alice 发送**正常的、合法的消息**，称为**正常文本（covertext）**，记作 `C`。
- 正常文本按照已知分布 P_C 生成，Eve 也掌握该分布。
- 可想象成正常文本由 Alice 可以访问的某个数据源生成。

2. Alice 处于活跃状态（包含隐写信息）

- Alice 发送**隐写文本（stegotext）**，记作 `S`，其分布为 P_S。
- 隐写文本是通过**嵌入函数** `F` 计算得到的，并包含要传递给 Bob 的**嵌入消息** `E`。
- **嵌入算法**可以使用：
- **私有随机源** `R`（仅 Alice 知道，与 `E` 和 `C` 独立）。
- **共享密钥** `K`（Alice 和 Bob 共享，Eve 不知道）。
- `K` 在使用隐写系统之前**通过安全信道随机选取并共享**，确保其与 `E`、`R` 和 `C` **相互独立**。

▼ 隐写系统模型

隐写系统的工作流程如下，Alice 端的**开关**（switch）决定她是处于活跃状态还是非活跃状态：

1. 开关在位置 0（Alice 非活跃）

- Alice 仅发送正常文本 `C` 到 Bob。
- 这些数据由正常文本数据源生成，不包含任何隐藏信息。
- Eve 监听到 `C`，但不会发现任何异常。

2. 开关在位置 1（Alice 活跃）

- Alice 需要发送消息 `E` 给 Bob。
- 她使用嵌入函数 `F` **将 秘密消息** `E` **隐藏到正常文本** `C` **中**，生成隐写文本 `S`：

```
S = F(C, K, R, E)
```

- 隐写文本 `S` 通过公共信道发送给 Bob。
- **Eve 和 Bob 都能观察到** `S`，但 Eve 不知道是否包含隐藏信息。
- Bob 使用**提取算法** `G` 以及共享密钥 `K` **从** `S` **中恢复消息**：

```
Ê = G(S, K)
```

- 目标是让 `Ê` 尽可能接近 `E`，即 Bob 能够正确接收信息。

▼ 隐写系统的安全性分析

从 Eve 的角度来看，她无法知道 Alice 是否处于活跃状态：

- **如果 Alice 处于非活跃状态**，Eve 观察到的数据分布是 P_C。
- **如果 Alice 处于活跃状态**，Eve 观察到的数据分布是 P_S。

Eve 的任务是进行**假设检验**（hypothesis testing）：

- 她需要区分她观察到的内容是 P_C（正常通信）还是 P_S（隐写通信）。
- **隐写系统的安全性取决于 P_C 和 P_S 的相似程度**。

▼ 安全性度量

- 使用相对熵（KL 散度）来衡量 `P_C` 和 `P_S` 之间的差异：

$$D(P_C \parallel P_S) = \sum P_C(x) log(P_C(x)/P_S(x))$$

- **定义 1（隐写系统的安全性）**：
  - **完美安全（Perfectly Secure）**：如果 $D(P_C \parallel P_S) = 0$，Eve **完全无法区分**两种情况，隐写信息是不可检测的。
  - **ε-安全（ε-Secure）**：如果 $D(P_C \parallel P_S) \leq \varepsilon$，Eve 只能在一定误差范围内进行检测。

▼ 隐写系统的额外说明

1. 完美安全隐写系统（Perfectly Secure Stegosystem）

- 在完美安全的隐写系统中，Eve **无法区分** P_C 和 P_S，因此**不会获取任何关于嵌入信息的知识**。
- 这类似于 Shannon 在密码学中的**完美保密（Perfect Secrecy）**。

2. 隐写系统必须是有用的

- **定义中的约束** `I(Ê; E) > 0` 确保 Bob **至少能够部分恢复消息** `E`。
- 这意味着隐写系统**不仅要隐藏信息，还要确保信息可以成功提取**。

3. 与其他隐写模型的区别

- 其他隐写模型通常假设 Alice **修改已有的正常文本**，例如修改图像文件中的像素来隐藏信息。
- 但在此模型中，隐写系统的目标是确保 P_C 和 P_S **在统计上完全相同**，即使 Eve **拥有原始正常文本**，她也无法检测到隐藏信息。

4. 完美安全与 Shannon 保密性

- 有人可能会认为，完美安全的隐写系统也应该满足 Shannon 定义的**完美保密性**（即隐写文本 `S` 和嵌入信息 `E` 统计上独立）。
- 但由于嵌入算法 `F` **不依赖于消息** `E` **的分布**，这已经保证了**隐写文本** `S` **对于 Eve 来说是独立于** `E` **的**，因此没有必要显式添加 Shannon 完美保密性的要求。

总结

- **隐写系统的目标**：让 Alice 能够向 Bob 发送隐藏信息，而 Eve 无法检测到信息的存在。
- **核心安全性分析**：通过比较 `P_C`（正常文本分布）和 `P_S`（隐写文本分布）的相似性(相对熵)来衡量隐写信息的可检测性。
- **完美安全隐写系统**：当 `D(P_C ‖ P_S) = 0` 时，Eve **无法区分**正常通信和隐写通信，隐写信息完全不可检测。

▼ 隐写系统：定义、安全性与评注

▼ 定义 1. 什么是隐写系统

固定一个原始载体文本分布 C (Covertext distribution) 和一个消息空间 E (Message space)

如果存在如上所述的随机变量 K (共享密钥) 和 R (随机性来源)，使得对于所有定义在 E 上且信息熵 H(E) > 0 的随机变量 E (待嵌入消息)，都满足 I(Ê; E) > 0 (Bob 解出的消息 Ê 与原始消息 E 之间的互信息大于0)，那么一对算法 (F, G)（F为嵌入算法，G为提取算法）被称为一个**隐写系统 (stegosystem)**。

此外：

- 如果满足 $D(P_C \| P_S) = 0$，原始载体分布 P_C 与隐写后载体分布 P_S 之间的相对熵为0，则该隐写系统被称为**完全安全 (perfectly secure)**（针对被动攻击者）。
- 如果满足 $D(P_C \| P_S) \leq \epsilon$，则该隐写系统被称为 ϵ**-安全**（针对被动攻击者）。

该模型描述的是一个**一次性使用 (one-time use)** 的隐写系统，其中爱丽丝要么一直处于活动状态（发送隐写信息），要么一直不处于活动状态。

如果爱丽丝向鲍勃发送多个相关的消息，并且其中至少一条包含隐藏信息，则认为她始终处于活动状态，此时 S 由她所有消息的串联组成。

> 1. 在一个完全安全的隐写系统中，Eve 无法区分两种分布（原始载体和隐写后载体），因此完全无法获知嵌入消息的存在性。这与香农 (Shannon) 为密码系统提出的完美保密 (perfect secrecy) 概念相似
> 2. 隐写系统定义中的条件 I(Ê; E) > 0 意味着该隐写系统是“有用的”，即鲍勃至少能获取到关于消息 E 的一些信息。
> 3. 我们的模型不同于有时为隐写术考虑的另一种场景，即Alice使用一个Eve已知的原始载体文本，并对其进行修改以嵌入隐藏信息。这类方案只能为那些对比修改后的隐写文本与原始载体文本能力有限的攻击者提供保护（否则，它们很容易被攻破）。例如，这适用于视觉图像上隐写术的流行用法，其中隐写图像对于人类来说可能在感知上与原始图像无法区分，但对于能够访问原始图像的算法来说则不然。

4. 在定义一个“完全安全”(指隐写层面，即$D(P_C \| P_S) = 0$）的隐写系统时，**无需额外显式地要求“嵌入的消息 E 对攻击者 Eve 也必须达到完美保密（即 S 和 E 统计独立)”**。

    a. **前提1：嵌入算法 F 的通用性。**
    模型要求嵌入算法 F 在设计时不知道（也不能依赖）待嵌入消息 E 的具体概率分布
    P_E。F 必须对 *任何* 可能的消息分布P_E 都能工作。

    b. **前提2：隐写系统的“完全安全”定义。**
    一个完全安全的隐写系统意味着，嵌入秘密信息后的载体 S 的概率分布
    P_S必须与原始无秘密信息的载体 C 的概率分布P_C完全相同，即它们的KL散度为零: $D(P_C \| P_S) = 0$

    c. **推论：P_S 的不变性。**
    结合前提1和前提2：由于嵌入算法 F 对于 E 的 *任何* 可能分布都必须产生一个遵循固定分布 P_C 的输出 S（因为 P_S必须等于 P_C 才能保证完全安全)，这意味着**无论具体嵌入的是哪个消息值 e，或者说无论消息 E 遵循何种先验分布 P_E，最终输出的隐写载体 S 的整体统计分布 P_S都必须保持不变，并且恒等于P_C。**

    d. **核心结论：S 与 E 对 Eve 而言统计独立。**
    如果对于所有可能的消息 E，产生的 S 的分布P_S都是同一个固定的分布（即 P_C)，那么从攻击者 Eve 的视角来看，她观察到的 S 的统计特性与具体是哪个 E 被嵌入是无关的。换句话说，S 的产生在统计上独立于 E。Eve 无法通过分析 S 的分布来获取任何关于 E 的信息，因为P_S始终是她已知的 P_C，不随 E 变化。

e. **最终结果：E 对 Eve 的完美保密是自动满足的。**
既然 S 与 E 对 Eve 而言是统计独立的，这就意味着消息 E 对于 Eve 达到了香农意义上的完美保密，而这并不需要作为一条额外的约束条件加入到定义中。

5. **如果隐写后生成的文本 S 的概率分布P_S 与原始（无辜的）载体文本 C 的概率分布 P_C 完全相同（即它们的相对熵 $D(P_C \| P_S) = 0$，或者说 $P_C = P_S$），那么这个隐写系统就是“完全安全的”。**

这里的核心在于**概率分布的相同性**，而不是仅仅“空间”相同。

- “文本空间”可以理解为所有可能的文本（或载体）的集合。原始文本和隐写后生成的文本可能都属于同一个文本空间。
- 但关键在于，在这个空间中，**各种文本出现的概率**是否因为嵌入了信息而发生了改变。

作者的逻辑是：

1. **攻击者 Eve 的视角**：Eve 观察公共信道上的数据。她知道原始无辜消息遵循分布 P_C。
2. **隐写操作**：Alice 嵌入信息 E，生成隐写文本 S，S 遵循分布 P_S。
3. **完全安全的定义**：如果P_S 和 P_C这两个概率分布完全一样，那么 Eve 无论观察到什么数据，都无法从统计上判断这个数据是来自P_C（无辜消息）还是来自 P_S（隐写消息），因为这两个“来源”产生数据的统计规律完全一致。
4. **结果**：Eve 无法检测到隐写行为的存在。

核心思想是让隐写操作后的输出在统计特性上与原始载体的统计特性无法区分。

6. 利用Oracle的无密钥隐写系统为何不算完美安全

   在我们对隐写系统的定义中，鲍勃通过一个“解码机(oracle)”得知爱丽丝是否处于活动状态。

   1. 人们可能会试图构建以下这种“完美的”隐写系统，它利用这种知识来传输隐藏信息而无需使用共享密钥。
   2. 不失一般性地，考虑一个嵌入算法 F，它包含一个理想的信源编码器，该编码器设法将某个消息 E1 ⇒ 隐写文本 S1，S1 由独立且均匀随机的比特组成。
   3. 如果原始载体文本 C 是相同长度的独立且均匀随机的比特序列，那么这两种分布是相同的，Eve无法区分压缩后的消息和原始载体文本。
   4. 在这种情况下，鲍勃无需任何密钥即可获得 E1。他区分隐写文本和原始载体文本的优势完全来自于那个解码机，人们可能会因此得出结论，假设存在这样的解码机会导致出现微不足道的隐写系统。
   5. 然而，这个结论并不成立，因为根据定义1，所描述的隐写系统并非完全安全。回想一下，F 是确定性的，并且被要求对任何消息分布都有效，因此它也必须对某个熵严格小于 E1 的消息 E2 有效
   6. 例如，当Eve对消息有部分了解时。令 S2 = F(E2)。那么直观上很清楚，确定性的 F 不会输出足够多的随机比特，因此 C 和 S2 的分布是不同的。形式上，这可以通过两种方式展开消息和隐写文本之间的互信息来看出。由于编码器是确定性且完美的，我们从展开 I(E1; S1) 得到 H(S1) = H(E1)。同样的编

码器应用于 E2 也唯一地确定了 S2，因此从展开 I(E2; S2) 得到 H(S2) = H(E2) - H(E2|S2) ≤ H(E2)。因此，根据对 E2 的假设，H(S2) ≤ H(E2) < H(E1) = H(S1)，这意味着 S1 和 S2 的分布不同，这与隐写系统是完美的假设相矛盾。

7. **至少，依赖 "oracle" 的特定无密钥隐写系统设计，不能达到其定义的“完美安全”。**

   本文的整个模型和“完美安全”的定义是**基于共享密钥 K 构建的**。即使我们引入一个强大的 "oracle" 来帮助 Bob 区分 C 和 S，也不能轻易地构造出一个满足其“完美安全”的无密钥系统。

   为什么那个特定的无密钥系统不完美安全，关键在于：

   - **完美安全的定义要求普遍性**：嵌入算法 F 必须对 *任何* 消息 E 的分布都产生与原始载体 C 分布相同的隐写后载体 S 的分布（即 $P_S = P_C$ 对所有 E 都成立）。
   - **该无密钥系统的局限性**：它依赖于信源编码将消息 E1 压缩成特定形式（独立均匀随机比特）的 S1，使得P_{S1}恰好等于P_C。但当消息变为熵较低的 E2 时，同样的确定性编码器 F 会产生熵也较低的 S2，导致 $P_{S2} \neq P_C$。

   是否意味着所有无密钥隐写系统都不能完美安全?

   这篇论文并没有直接下这个普适性的结论。它主要关注的是在**其提出的信息论模型和定义框架下**的安全性。

   - **依赖定义的**：如果其他研究者提出不同的“完美安全”定义，或者不同的隐写系统模型（不一定依赖共享密

钥 K，或者对嵌入算法F的要求不同），那么结论可能会不一样。

- **理论可能性**：理论上，如果存在某种方法，即使没有预共享密钥，也能确保对于任何消息 E，产生的 P_S 都等于 P_C，那么按照这篇论文的安全性核心标准（$D(P_C \| P_S) = 0$），它也将被认为是“完美安全”的。但实现这种无密钥方案的难度非常大，尤其是要满足对“任何消息分布都有效”的普遍性要求。

总结来说：

1. 在这篇论文的框架下，其定义的“完美安全”是与共享密钥 K 紧密相关的。
2. 简单地用 "oracle" 替代密钥，并不能轻易构造出满足其“完美安全”定义的无密钥系统，因为其对嵌入算法F的通用性要求很严格。
3. 它并没有绝对地断言“所有无密钥隐写系统都不能完美安全”，但暗示了在其模型下，实现这一点是困难的，至少所举的例子失败了。

根据这篇论文的特定定义和模型，作者论证了他们所描述的那个依赖 "oracle" 的无密钥隐写方案无法达到“完美安全”，因为其嵌入算法不具备必要的通用性。

▼ 定义2. 随机过程

1. 通常将信息源建模为随机过程是合适的。例如，原始载体文本 (covertext) 可能由同一实验的独立重复生成。
2. 在上述模型中，Eve 观察完整的原始载体文本，但考虑一个受限的攻击者也具有意义，该攻击者只能访问长原始载体文本序列的一个子集。

让上述模型中的所有随机变量扩展到随机过程，并令 n 表示重复的次数。

假设原始载体文本是由一个**平稳信息源 (stationary information source)** 生成的。

因此，在夏娃被限制只能看到原始载体文本序列的有限部分的情况下，原始载体文本和隐写文本 (stegotext) 过程之间的**归一化相对熵 (normalized relative entropy)** 决定了安全性。

定义 2. 对于具有平稳原始载体文本的随机过程的隐写系统，如果满足以下条件，则称其为**平均完美安全 (perfectly secure on average)**（针对被动攻击者）：

$$\lim_{n\to\infty}\frac{1}{n}D(P_C\|P_S)=0$$

类似地，对于随机过程的隐写系统，如果满足以下条件，则称其为
平均ϵ-安全（针对被动攻击者）：
在整个实验过程中，爱丽丝仍然要么一直处于活动状态，要么一直处于非活动状态，并且通常情况下，隐写文本的分布不具备遍历性 (ergodic)

$$\lim_{n\to\infty}\frac{1}{n}D(P_C\|P_S)\leq\epsilon.$$

解释：

这句话是对定义2中将模型扩展到随机过程的几点重要澄清和强调：

1. **"Alice is still either active or inactive during the entire experiment"**
   - **含义**：即使我们将原始载体和隐写载体视为由一个随机过程生成的长序列（由 \(n\) 次重复组成），爱丽丝的“模式”选择（是发送无辜的原始载体 C，还是发送嵌入了信息的隐写载体 S）是在整个序列（整个实验）开始之前就决定的，并且在整个序列的传输过程中保持不变。
   - **重要性**：这意味着我们不是在序列的每个时间点或每个子段独立决定爱丽丝是否活动。要么整个长序列都是 C，要么整个长序列都是 S。这简化了分析模型，因为攻击者 Eve 面对的是两个固定的、整体的假设之一（整个序列是 C 或整个序列是 S）。
2. **"the stegotext distribution will not be ergodic in general"**
   - **背景知识：遍历性 (Ergodicity)**
     - 在随机过程中，如果一个过程是遍历的，那么其长时间的**时间平均 (time average)** 等于其**统计平均 (ensemble average 或**

expectation)。简单来说，通过观察单个足够长的样本路径，就可以推断出整个过程的统计特性。

- 对于平稳过程，遍历性是一个更强的条件。平稳性只保证统计特性不随时间推移而改变，而遍历性则意味着单个实现可以代表整体。

- **含义**：作者指出，即使原始载体文本 C 假定来自一个平稳信息源（甚至可能具有遍历性），当爱丽丝嵌入信息 E 生成隐写文本序列 S 时，S 作为一个随机过程，**通常情况下不会是遍历的**。
- **原因（推测，基于隐写操作的性质）**：
  - **嵌入操作的确定性部分**：嵌入消息 E 本身通常是一个固定的秘密，或者其随机性与载体的随机过程不同步。嵌入算法 F，即使加入了随机性 R 和密钥 K，对于一个给定的消息 E，其对原始载体 C 的修改方式在某种程度上是系统性的，而不是完全遵循载体 C 的那种纯粹的、内生的随机演化。
  - **密钥 K 的固定性**：共享密钥 K 在整个过程中是固定的。
  - **消息 E 的固定性或有限性**：要嵌入的消息 E 通常来自一个有限的消息空间，或者在一次通信中是固定的。
  - 这些因素导致隐写文本序列 S 的统计特性可能无法仅通过观察其长时间行为来完全代表其“整体平均”特性，因为它受到了外部（消息E、密钥K）和确定性（嵌入算法F）因素的持续影响。其时间平均可能不等于（或者说不收敛于）一个假想的、仅由 P_S 定义的系综平均，特别是如果 P_S 本身就因为嵌入了特定信息而变得“不自然”或具有某种内在模式。
- **重要性**：
  - 这提示了分析隐写文本 S 的统计特性时可能遇到的复杂性。不能简单地假设可以通过观察一段长S序列就能完全掌握其 underlying 的概率分布 P_S 的所有特性。
  - 它也可能暗示了某些基于时间平均的检测统计量可能不适用于非遍历的隐写过程。
  - 尽管定义2中使用了 $D(P_C \| P_S)$，这里的 P_S 指的是整个隐写文本序列的联合概率分布，而非简单地假设 S 也是一个“良好行为”的遍历过程。

总结来说：

这段补充说明提醒读者，即使模型扩展到了随机过程，一些基础假设仍然保持（爱丽丝的活动状态在整个实验中是固定的），并且引入了一个关于隐写文本过程统计特性的重要告诫（它通常不是遍历的）。这对于如何正确理解和应用定义2中的安全性概念至关重要。归一化相对熵 $\frac{1}{n} D(P_C \| P_S)$ 仍然是核心度量，但 P_S 本身的复杂性（非遍历性）值得注意。