Data-Free Stochastic Attractors

A Whole New Paradigm in Neural Networks

James J Kalafus

September 10, 2024

Abstract

Traditional neural networks rely on data-driven training, leading to deterministic attractors that make outputs predictable. This "white paper" introduces **Data-Free Stochastic Attractors**, a paradigm-shifting approach that removes the reliance on training data, focusing solely on *metrics* such as *stochasticity* and *entropy*. This method opens the door for applications in **cryptography** and **obfuscation**, offering a fundamentally different way of thinking about neural networks.

A. Introduction

We propose a new framework: **Data-Free Stochastic Attractors**. In this approach, the network's behavior is determined without training data, relying solely on the **loss metric** to optimize for key metrics, such as:

- **Stochasticity** — Measure how sensitive output is to perturbations in the input.
- **Entropy** — Quantifies system information content, uncertainty, and randomness.
- **Stability** — Prevent numerical instability during optimization.

By removing the constraints imposed by reference data, this paradigm is suited for applications such as **cryptography** and **obfuscation**, where unpredictability, randomness, and sensitivity to input variations are essential.

In conventional neural networks, *training data* plays a central role. Networks are optimized to converge on specific patterns or attractors within the data, making them well-suited for tasks such as classification or regression. However, this reliance on data also leads to *deterministic attractors*, where network behavior becomes predictable, which is ineffective for tasks requiring *unpredictability* and *stochasticity*, such as *encryption*.

Data-Free Stochastic Attractors break from this tradition by focusing on **unpredictability**, making network behavior ideal for applications where stochastic responses to input variations are crucial.

Deterministic Attractor: A system state that is insensitive to initial conditions. Similar inputs (constraints) lead to similar, stable outcomes (convergent solutions).
Stochastic Attractor: A system state that is sensitive to initial conditions. Small variations in input (constraints) lead to dissimilar, unpredictable outcomes (divergent solutions).
Stochasticity: Degree of sensitivity of outputs to small changes in inputs.

B. The Problem with Deterministic Attractors

Deterministic Behavior in Data-Driven Models

Neural networks trained on fixed datasets tend to converge on *deterministic attractors*, resulting in predictable outputs for given inputs. This is beneficial for data-fitting tasks but presents challenges when *unpredictability* or *stochastic behavior* is required. In fields like cryptography, where unpredictability is a core requirement, deterministic attractors pose a risk.

Removing Training Data

By eliminating the training dataset, the neural network is no longer bound by deterministic attractors. Instead, the network's behavior is driven purely by the properties optimized in the loss function, freeing it from the specific patterns present in any dataset. Building the loss function around *stochasticity and entropy metrics* makes networks highly sensitive to small variations in input, leading to *stochastic attractors*.

C. Why This Is a Groundbreaking Approach

Data-Free Networks

In most neural network applications, models are trained using supervised learning techniques where the network fits to a dataset. In contrast, the **Data-Free Stochastic Attractor** model focuses purely on *network dynamics* without requiring any reference data. This unleashes high *stochasticity* and *entropy*, resulting in high variance output.

Stochastic Attractors in Neural Networks

In conventional neural networks, models are trained to minimize error over a specific dataset, which results in convergence to **deterministic attractors**—stable points or patterns that the model learns to predict consistently based on the data.

This is suitable for tasks like classification and regression but problematic for applications requiring **unpredictability**.

In the **Data-Free Stochastic Attractors** paradigm, the network is not constrained by fitting to training data. Instead, the **loss function optimization** becomes the primary driver, focusing on objectives such as **maximizing stochasticity** and **entropy**. The loss function penalizes predictable, deterministic behavior and promotes sensitivity to input changes, pushing the model to explore **stochastic attractors**—unstable points where even slight variations in input can lead to dramatically different outputs. This behavior is highly desirable for tasks like encryption, where unpredictability is key.

By designing the loss function to prioritize **stochasticity**, **entropy**, and **stability**, the model evolves into a system that resists converging to deterministic attractors. Instead, it hovers in zones of high sensitivity, where the output is shaped by subtle nuances in the input. This unpredictability is particularly valuable in cryptographic systems, which benefit from **high variance** and **difficulty in reverse-engineering** the transformations.

D. Applications

Cryptography

The intrinsic *high variance* of *stochastic attractors* make this approach ideal for cryptography. In traditional encryption, the security of asymmetric key systems relies on the difficulty of reversing the encryption without the correct key and the challenge of reverse engineering the key itself. *Stochastic attractors* offer a distinct advantage: their sensitivity to small input perturbations results in *significant, unpredictable changes* in the output. This makes it intractable to train or derive an encryption key clone from the reverse image of the decryption key, effectively securing the cryptographic process against potential attacks that exploit key reversal.

Obfuscation

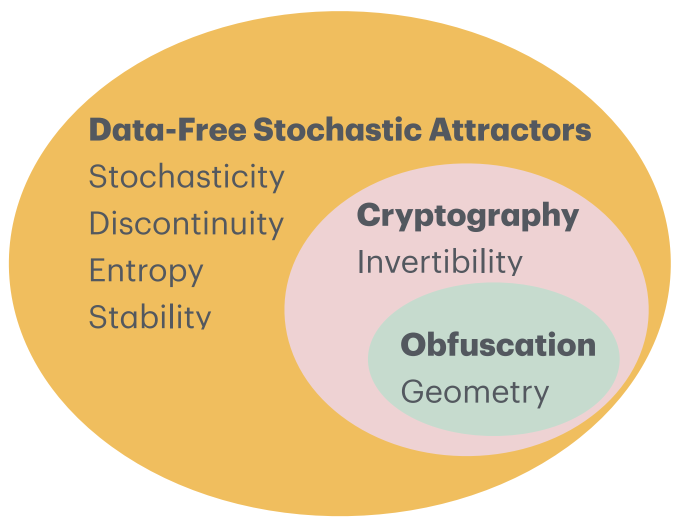
In fields like *privacy-preserving machine learning, homomorphic encryption, code obfuscation,* and *secure data sharing*, this approach can make data obfuscation more effective by generating stochastic encodings that are difficult to reverse-engineer without the correct attractor key, similar to cryptographic encodings but preserving critical structures and relationships of data and code.

E. Design and Implementation

Loss Function Design

The core of this technique is the design of the loss function, which optimizes for *stochasticity*, *entropy*, *stability*, *geometry*, and *invertibility*. The proposed loss function is:

$$\begin{aligned}\mathscr{L}_{\text{stochastic\_attractors}} = \\ &\lambda_{\text{stochasticity}} \cdot \mathscr{L}_{\text{stochasticity}} \\ &+\lambda_{\text{entropy}} \cdot \mathscr{L}_{\text{entropy}} \\ &+\lambda_{\text{discontinuity}} \cdot \mathscr{L}_{\text{discontinuity}} \\ &+\lambda_{\text{stability}} \cdot \mathscr{L}_{\text{stability}} \\ &+\lambda_{\text{geometry}} \cdot \mathscr{L}_{\text{geometry}} \\ &+\lambda_{\text{invertibility}} \cdot \mathscr{L}_{\text{invertibility}}\end{aligned}$$



Explanation of Loss Metrics

The design of the *loss function* determines the specific behaviors and applications of the model. Different metrics such as *stochasticity*, *entropy*, *discontinuity*, *stability*, *invertibility*, and *geometry* guide the model toward a range of potential uses, such as *cryptography* or *obfuscation*.

- **Cryptography**: Requires invertibility to ensure that data can be accurately recovered from an encoded state, without needing to retain geometric structure. This makes *invertibility* the key metric in cryptographic applications, where data must be fully reconstructible.
- **Obfuscation**: In addition to invertibility, geometry is included to preserve internal structures and relationships. This ensures that while the data is transformed into a less recognizable form, it remains in a usable structure for tasks like privacy-preserving machine learning or code obfuscation.

A visualization of these choices can be depicted as nested rings (see previous pg):

- The outer ring (generic **data-free stochastic attractor** design) includes key elements: *stochasticity*, *entropy*, *discontinuity*, and *stability*.
- The middle ring represents **cryptography**, which adds *invertibility* to the outer ring elements, ensuring data can be recovered.
- The inner ring represents **obfuscation**, adding *geometry* to retain structural integrity while obfuscating.

By choosing different combinations of loss metrics, the network can be optimized for specific goals. For example:

- Invertibility without geometry leads to **cryptographic** applications, where the focus is on the full recovery of the original data.
- Invertibility with geometry leads to **obfuscation** applications, where the data is transformed but remains usable in its obfuscated state.

Regardless of the specific design goals, these choices in the loss function guide the behavior and characteristics of the model.

Maximizing Unpredictability of Stochastic Attractors

A key goal is to prevent the network from collapsing into deterministic attractors. The loss function drives the model toward maximizing *entropy* and *stochasticity*, ensuring that outputs are diverse and non-predictable, even in the absence of training data.

F. Future Potential

Broader Applications

While **Data-Free Stochastic Attractors** show clear promise in cryptography and obfuscation, their flexibility also opens doors to **creative applications** in **Generative AI**. By focusing on metrics such as **stochasticity** and **geometry**, and leaving invertibility aside, this framework can serve as a foundation for several novel generative techniques.

Generative AI

1. **Generative Art**: By maximizing stochasticity and geometry, models can generate *unpredictable, non-repeating patterns* in visual or audio data. This allows for unique creative outputs in applications such as art, music, and design.

2. **Procedural Generation**: Models built on stochastic attractors can create *virtually unlimited novel outputs* for dynamic content generation in areas like *gaming*, *music*, and *interactive media*. These models could generate everything from new musical motifs to landscapes and video game levels.

3. **Hybrid Semi-Stochastic Attractors**: Combining high stochasticity with *highly focused data sets* compatible with *reduced entropy* loss targets enables *impressionistic reconstructions* of data. This approach allows for *efficient storage* while retaining enough variation for creative, unique outputs.

4. **Structured Creativity with Low Entropy**: By intentionally keeping *entropy low*, these models can produce *structured variations* that adhere to specific themes or styles while maintaining high stochasticity for *creative diversity*. This opens up exciting possibilities for models that balance randomness and predictability in artistic endeavors.

Building a Community

This approach is highly *generalizable* and can be adapted by researchers in fields such as cryptography, AI, and data processing. By sharing the technique in open forums and conferences, this paradigm could catalyze a range of innovations across industries.

G. Conclusion

The **Data-Free Stochastic Attractor** paradigm redefines how we think about neural networks by shifting from data dependency to a *loss-driven*, *stochastic* model. This approach creates unprecedented opportunities in fields where **unpredictability**, **sensitivity to inputs**, and **randomness** are paramount, particularly in **encryption** and **obfuscation**. By removing the need for training data, we unlock neural network behavior that is driven by stochasticity, entropy, and structure preservation, offering novel methods of encryption and **data security**.

This proposal represents an **urgent opportunity** for the research community to **break new ground** in cryptography—an area that has long sought machine learning solutions to tackle its most complex problems. The potential for **stochastic neural networks** in these areas remains vastly untapped, and with the techniques outlined here, the door is open for profound innovation.

Next Steps

- **Proof of Concept**: The next critical step is to develop a **proof of concept** encryption algorithm based on this paradigm. This would validate the theory, transform it into a practical solution, and firmly establish **Data-Free Stochastic Attractors** as a new, reliable elementary branch of machine learning.

    - Potential challenges include conservation laws surviving stochastic attraction, but these are likely **mitigable** and can be solved through careful architectural design.
    - **Timeline**: Initial proof of concept should be achievable within the next 6 to 12 months, with clear milestones such as developing a functional encryption prototype, followed by rigorous testing.
- **Collaborative Opportunities**: Given the broad applications of this technique, there is immense potential for **cross-disciplinary collaborations** in cryptography, security, data processing, and machine learning. By working together, researchers and practitioners can refine, implement, and expand the scope of **Data-Free Stochastic Attractors**, bringing this innovative approach into practical use.

This is a **call to action** for researchers and developers to embrace the potential of this paradigm and explore its applications beyond what has been traditionally possible. The road ahead is filled with opportunities to **reshape cryptography** and beyond with data-free stochastic approaches, paving the way for a new generation of secure, dynamic, and highly adaptive systems.

Appendix

Stochasticity — Measure how much perturbations of input change the output.

- $\mathcal{L}_{\text{sensitivity}} = \mathbb{E}_{x \sim p(x)} \left[\left\| J_f(x) \right\|_F^2 \right]$
- $\mathcal{L}_{\text{stability}} = \mathbb{E}_{x \sim p(x)} \left[\left\| \nabla_x f(x) \right\|_F^2 \right]$
- $\mathcal{L}_{\text{curvature}} = \mathbb{E}_{x \sim p(x)} \left[\left\| H_f(x) \right\|_F^2 \right]$

Entropy — Measure information content, uncertainty, and randomness.

- $\mathcal{L}_{\text{shannon\_entropy}} = \mathbb{E}_{x \sim p(x)} \left[-\sum p(\theta_i) \cdot \ln p(\theta_i) \right]$
- $\mathcal{L}_{\text{mutual\_information}} = \mathbb{E}_{x \sim p(x)} \left[\log p(x) \right] - \mathbb{E}_{x \sim p(x), \theta \sim p(\theta|x)} \left[\log p(x \mid \theta) \right]$

Discontinuity — Prevent numerical instability.

- $\mathcal{L}_{\text{norm\_discontinuity}} = \mathbb{E}x \sim p(x) \left[\max \left(0, \left| f(x+\delta) - f(x) \right| - \epsilon_{\text{threshold}} \right) \right]$
- $\mathcal{L}_{\text{directional\_discontinuity}} = \mathbb{E}x \sim p(x) \left[\max \left(0, \left\| \frac{\partial f(x)}{\partial \mathbf{v}} - \frac{\partial f(x+\delta)}{\partial \mathbf{v}} \right\| - \epsilon_{\text{threshold}} \right) \right]$
- $\mathcal{L}_{\text{gradient\_discontinuity}} = \mathbb{E}x \sim p(x) \left[\left\| \nabla_x f(x) - \nabla_x f(x+\delta) \right\|_F \right]$
- $\mathcal{L}_{\text{hessian\_discontinuity}} = \mathbb{E}x \sim p(x) \left[\left\| \nabla_x^2 f(x) - \nabla_x^2 f(x+\delta) \right\|_F \right]$
- $\mathcal{L}_{\text{lipschitz}} = \mathbb{E}x \sim p(x) \left[\sup_{y \neq x} \frac{\left\| f(x) - f(y) \right\|}{\left\| x - y \right\|} \right]$

Stability — Prevent numerical instability.

- $\lambda_{\text{adaptive\_regularization\_coefficient}} = \lambda_0 \cdot \left(1 + \mathbb{E}_{x \sim p(x)} \left[\log p(x) \right] - \mathbb{E}_{x \sim p(x), \theta \sim p(\theta|x)} \left[\log p(x \mid \theta) \right] \right)$
- $\mathcal{L}_{\text{overflow\_penalty}} = \mathbb{E}_{x \sim p(x)} \left[\sum \max \left(0, \left| f(x) \right| - \text{max\_threshold} \right) \right]$
- $\mathcal{L}_{\text{regularization}} = \mathbb{E}_{x \sim p(x)} \left[\ln \left(1 + \left| f(x) \right| \right) \right]$
- $\mathcal{L}_{\text{norm}} = \mathbb{E}_{x \sim p(x)} \left[\left| f(x) \right|^2 \right]$

Appendix, cont.

Reversibility — Measure information loss or conservation.

- $\mathscr{L}_{\text{invertible}} = \mathbb{E}_{x \sim p(x)} \left[\sum_{i,j} \max \left(0, \epsilon - \left| J_f(x)_{ij} \right| \right) \right]$
- $\mathscr{L}_{\text{orthogonal}} = \mathbb{E}_{x \sim p(x)} \left[\left\| J_f(x)^\top J_f(x) - I \right\|_F^2 \right]$
- $\mathscr{L}_{\text{redundancy}} = \mathbb{E}_{x \sim p(x)} \left[\left\| \text{Cov}\big(f(x)\big) - I \right\|_F^2 \right]$

Geometry — Measure internal structures and relationships.

- $\mathscr{L}_{\text{distance}} = \mathbb{E}_{(x_1, x_2) \sim p(x)} \left[\left(\left| f(x_1) - f(x_2) \right| - \left| x_1 - x_2 \right| \right)^2 \right]$
- $\mathscr{L}_{\text{volume}} = \mathbb{E}_{x \sim p(x)} \left[\left| \det J_f(x) - 1 \right| \right]$
- $\mathscr{L}_{\text{time\_consistency}} = \mathbb{E}_{(x_t, x_{t+1}) \sim p(x)} \left[\left| f(x_t) - f(x_{t+1}) \right|^2 \right]$