Applications of Chaotic Dynamical Systems in Cryptography

William Kennedy
Sean Lawlis
University of Ottawa

April 27, 2023

Abstract

In this project we explore the use of chaotic discrete dynamic systems as cryptographic methods as a possible alternative to the modern standard of algebraic, geometric, and number theoretic cryptology. This theorized approach has been considered many times over the past 30 years, with different chaotic discrete dynamic systems being considered as candidates for new encryption schemes[2],[3]. Specifically we will consider the implementation of chaotic discrete dynamics as symmetric key ciphers in (cf [1],[2],[3]).

1 Introduction

1.1 Background

In 1949 Claude Shannon published his seminal paper(cf [4]). His paper discussed the requirements of Secrecy Systems, that is, what is required to develop a secure cryptographic scheme. Claude defined two important ideas; confusion which refers to obscuring the relationship between the ciphertext and the secret key as much as possible, and diffusion which refers to deteriorating the statistical structure of the plaintext and the ciphertext. These requirements are integral when developing cryptographic systems, and it is theorized that chaotic dynamic systems can embody these requirements. In the literature of chaotic cryptography the accepted definition is the one defined by Matthew Devaney where a dynamic system is chaotic if ; there is a sensitive dependance on initial conditions; the dynamic map is topologically mixing, and the periodic points are dense.

The traditional approach of developing cryptographic schemes has been based in algebraic, geometric, and number theoretic ideas. Some examples of which use linear transformations of vectors, generators of cyclic groups, or isomorphisms between quotient rings and subspaces to develop different ways to encrypt, decrypt, and correct errors when creating encryption schemes. This project considers the novel approach of using chaotic dynamic systems as an encryption scheme.

1.2 Chaos

Chaos theory is the study of dynamical systems that have unpredictable trajectories and the long-term behaviour cannot be predicted.

Many definitions of chaos exist, but in the literature of Chaotic Cryptology Robert Devaney's definition of chaos is the widely accepted version. A discrete dynamical system of the form

$$x_{k+1} = f(x_k), \quad x_0 \in I,$$

is chaotic if it satisfies the following properties:

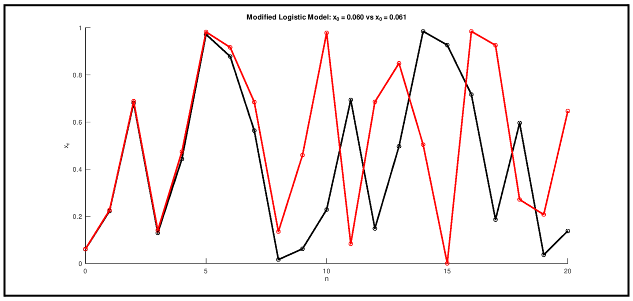
i. It has sensitive dependence on initial conditions
ii. It is topologically mixing
iii. Its periodic points are dense

i. Sensitive dependence on initial conditions

$$\exists \delta > 0\ \forall x_0 \in I, \varepsilon > 0\ \exists n \in \mathbf{N}, y_0 \in I : |x_0 - y_0| < \varepsilon$$
$$\Rightarrow | f^n(x_0) - f^n(y_0) | > \delta$$

What this mathematical definition says: Small differences in the initial conditions yield very different trajectories, which makes long term behaviour impossible to predict. In cryptography, this is an ideal characteristic to have because if you only know the outcome, it's computationally infeasible to go backwards and find the starting state. Even a moderately accurate approximation of the initial condition will give the incorrect trajectory due to the sensitive dependence. The initial condition will act as the secret key, its trajectory (ciphertext in this context) will be highly obscured even if there is a slight deviation in the initial conditions. This is important for satisfying the confusion principle defined by Claude Shannon.

Comparing trajectories of two similar initial conditions



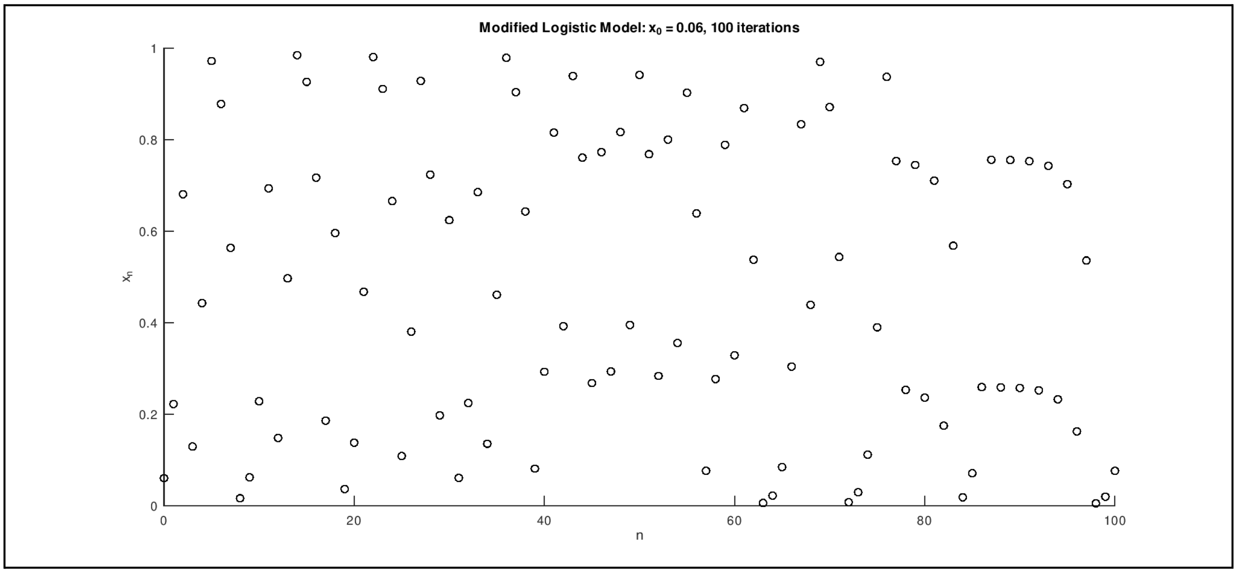
This graph shows the divergent trajectories of two very close initial values, 0.060 and 0.061, over 20 iterations. This demonstrates the principle of sensitive dependence on initial conditions.

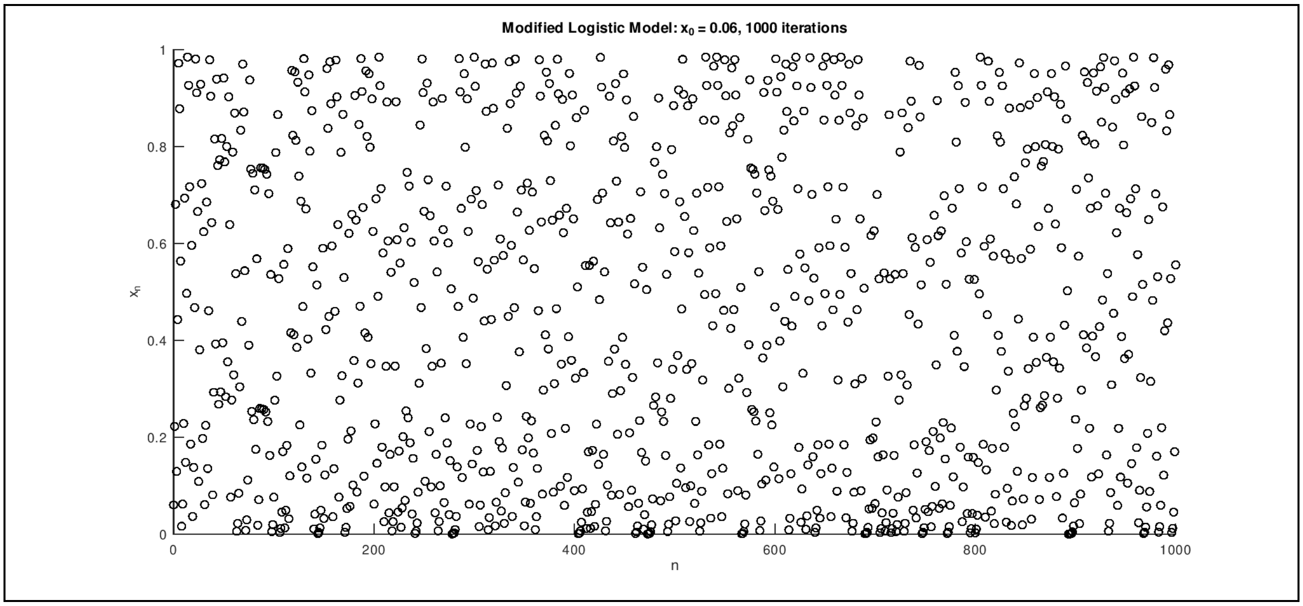
ii. Topological mixing

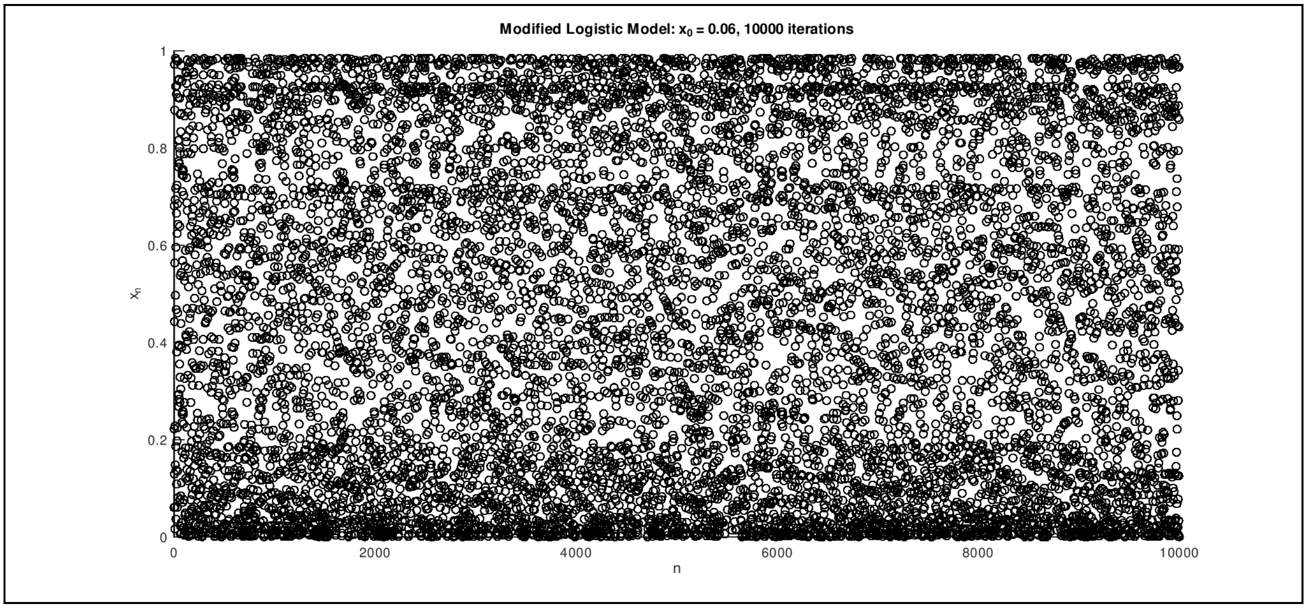
When a subset of the domain eventually intersects every possible open subset of the range of the discrete iteration function, for an open subset A of the domain and open subset B of the range we have:

$$B \cap f^n(A) \neq \emptyset$$

No matter where the initial conditions start, they will eventually intersect with every open subset of the range. Take a small open subset, then the range of that map will intersect with every possible subset of the codomain after n iterations. At some point in time two trajectories of any two initial conditions will become arbitrarily close to each other. Which implies there is a high degree of randomness and unpredictability, meaning that nearby points can diverge quickly and become arbitrarily separate. In the context of cryptography this means that every possible key will eventually be mapped to every possible output, thus diffusing the statistical relationship between the secret key and the ciphertext.







The first graph shows 100 iterations of the modified logistic map (MLM). We can imagine an interval such that none of the points in that interval have been mapped to by the MLM after 100 iterations. Topological mixing guarantees that we will eventually map to a point in that interval after some number of iterations (no matter how small the interval).

iii. Dense periodic points

For the phase space to have dense periodic points, each point in the range must be approached arbitrarily close by periodic orbits. This complements the system's sensitive dependence on initial conditions, ensuring that perturbations in any initial condition will always lead to a different periodic orbit..

This is a meaningful property in the context of cryptography, since being in a cyclic orbit does not compromise the initial key and no given outputs of the key are compromised since infinitely many keys can map to the same point.

1.3 Cryptography- Stream Cipher

In algebra we view the basic encryption of a stream cipher as vector addition. A stream cipher is a process where we encrypt a message letter by letter, or more specifically, coordinate by coordinate. Our message and secret key are represented as vectors, their addition will yield our cipher text. This idea is illustrated by the following example:

- Alice and Bob want to communicate securely over an unsecure network.
- **Setup**: A secret key is predetermined and shared between the sender Alice, and the recipient Bob
- **Goal**: Encrypt and decrypt the message "HELLO" using a secret key
- Each coordinates work in modulus 26:
  - (H, E, L, L, O) + (X, L, E, D, V) = (F, Q, Q, P, K)
  - (8, 5, 12, 12, 15) + (24, 12, 5, 4, 22) = (6, 17, 17, 16, 11)
- **Result**: a word that has no relationship to our plaintext word and you cannot work backwards to find the plaintext without the secret key.

---

2 Models

2.1 Logistic Map

The logistic map is a model that we have seen in class, and it serves as a great entry-point into the study of chaotic dynamical systems.

$$x_{n+1} = rx_n(1 - x_n)$$

r is typically interpreted as the “reproduction rate.” For different values of r we have different behaviour. For r values between 0 and 3.57 we have convergence or permanent cycles. For r between 3.57 and 4 we see chaotic behaviour – this is what we are interested in. For r > 4 the logistic map is unbounded (it blows up to infinity).

x_n typically represents the ratio of current population to maximum possible population. Thus it lies between 0 and 1 (i.e. the image of x_n is the open unit interval). Why use the Logistic map?

1. Relative simplicity: the logistic map is a simple polynomial of degree 2.
2. Computationally efficient: computers can compute many iterations of this function easily. Later, we will introduce a different model that involves exponentiation, which is noticeably slower to compute.
3. Well-studied: the logistic map has been extensively studied.

2.2 Modified Logistic Map (MLM)

The MLM allows for a larger chaotic range for the parameter r; the logistic map has a chaotic range of [3.6,4] for the parameter r. In the MLM the chaotic range for the parameter r is [2,4], this is a five-fold increase allowing for stronger security since r can assume more values and an increased sensitive dependence on initial conditions.

$$X_{n+1} = \begin{cases} g(x) = rX_n(1 - X_n), & X_n < 0.5 \\ h(x) = rX_n(X_n - 1) + \frac{r}{4}, & X_n \geq 0.5 \end{cases}$$

Below are figures illustrating the chaotic range, trajectories of the return map, and the trajectory of a random value within the unit interval.

MLM: Bifurcation Diagram

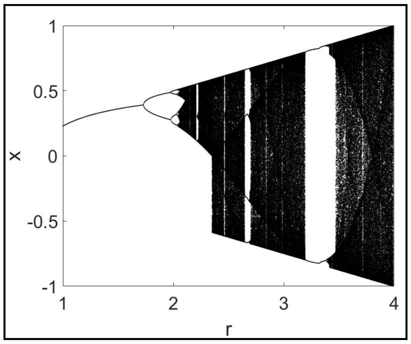
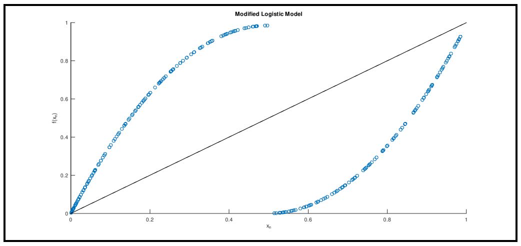


Figure 1: Looking at this bifurcation diagram we can see that the MLM is chaotic for r values between 2 and 4. We can also see that there are sections or “islands” of stability. We don’t consider these r values.

MLM: Return Map



MLM: Trajectory

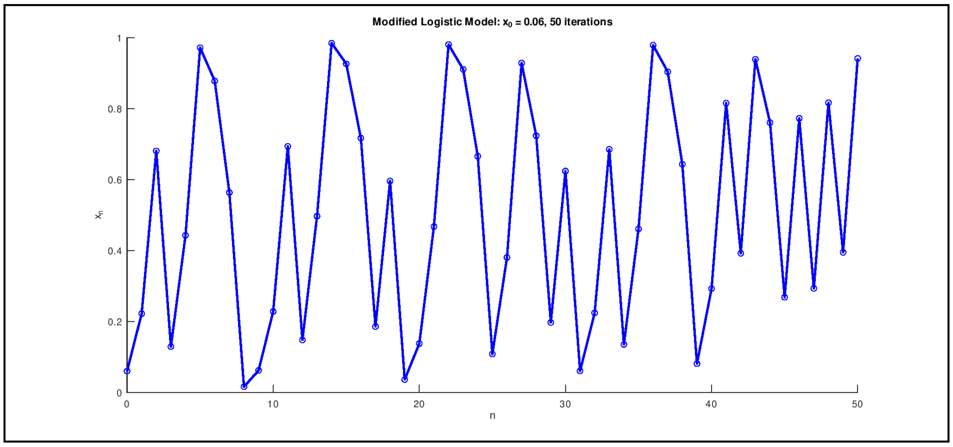
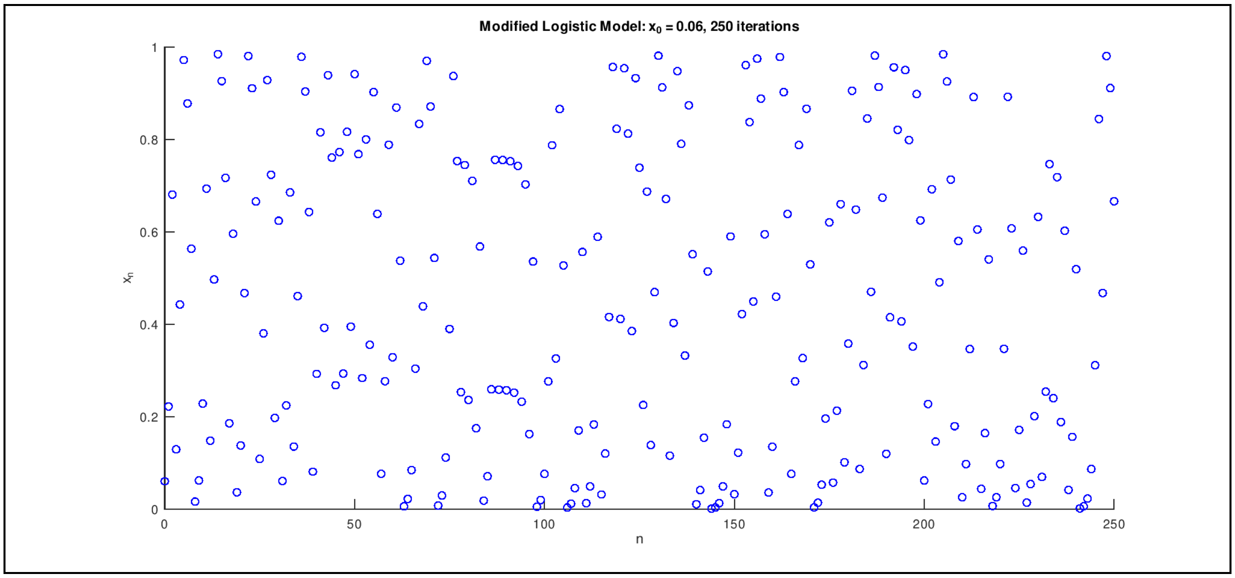


Figure 3: graph shows the trajectory of the MLM, starting with an initial value of 0.06 (arbitrary), over 50 iterations. The x-axis is the number of the iterations, and the y-axis shows the values of each iterate. Its chaotic trajectory resembles randomness, which is exactly what we want. This graph effectively shows us what the MLM is doing as we repeatedly iterate. Note that for any initial value x0, we will always get the exact same trajectory. The trajectory of the MLM is completely deterministic! This is essential for its use in cryptography. It is also easy to calculate in one direction, while not in the other.

MLM: Chaotic Distribution



This graph depicts 250 iterations. Once again, the distribution appears random.

2.3 Ricker Model

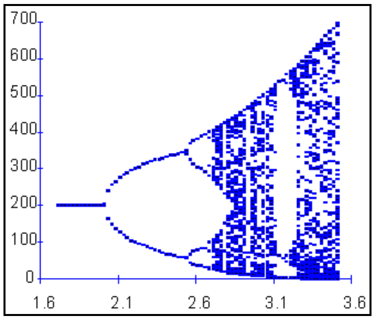
$$N_{t+1} = N_t e^{r\left(1-\frac{N_t}{k}\right)}$$

The Ricker model was introduced in 1954 by Bill Ricker in the context of stock and recruitment in Canadian fisheries. Our use of this chaotic discrete dynamical system in cryptography is a novel extension of existing research. Typically, N_t is the number of individuals in the population at generation *t*, *r* is the intrinsic growth rate, and *k* is the carrying capacity. The Ricker model satisfies the 3 properties of chaos outlined earlier. For consistency, we will write the Ricker model with *x* instead of *N*.

$$x_{n+1} = x_n \cdot e^{r \cdot (1 - x_n / k)}$$

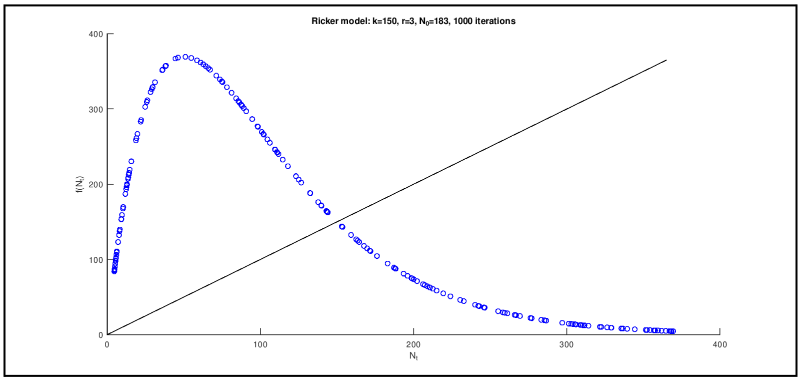
The parameter r functions as a key as usual, and *k* functions as an additional key (this extra key is an advantage of this model). The chaoticity depends on the chosen value of r: We used the chaotic range of r in [2.9,3] and k in [100, 300].

The bifurcation diagram

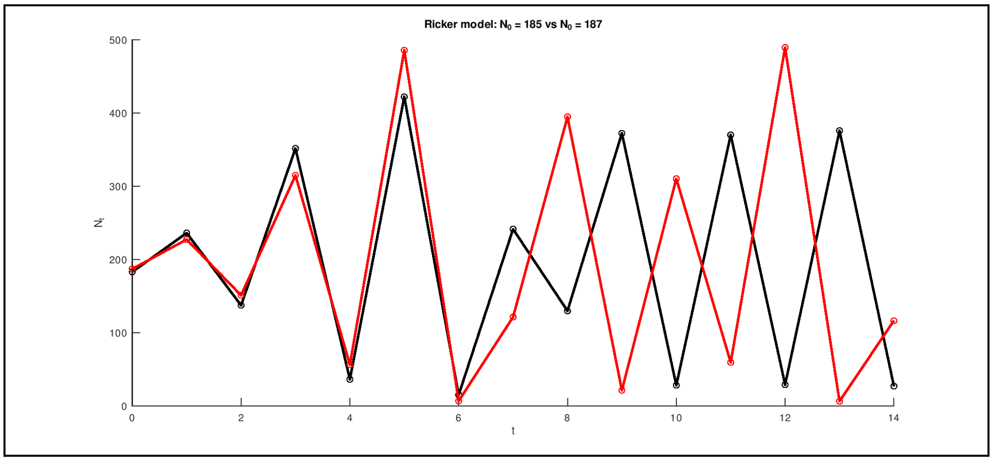


*Shows us the values of r that make the Ricker model chaotic (provide source)******

Ricker Model: Return Map

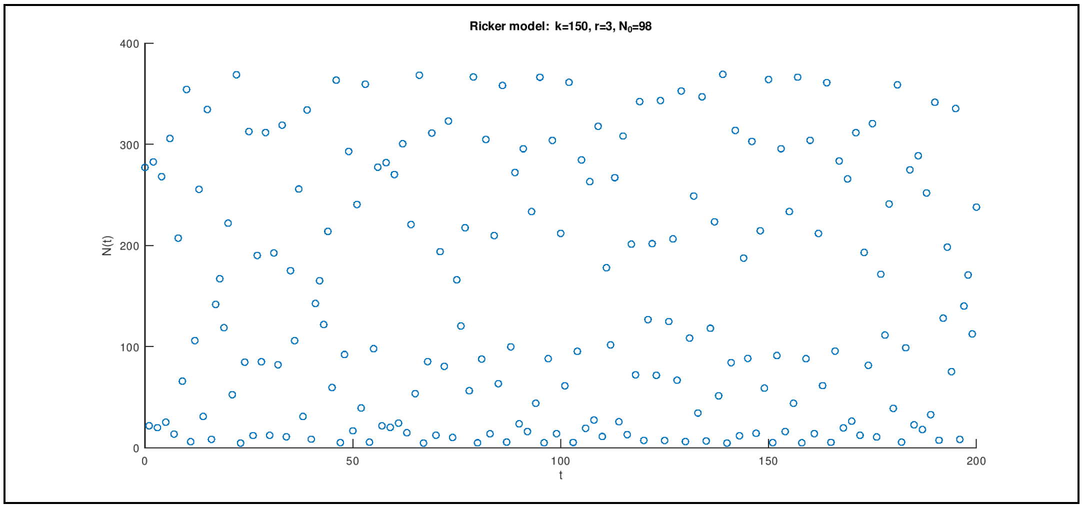


Ricker Model: Trajectory



Just like with the MLM, we can see that the Ricker model is highly sensitive to initial conditions. We have two very close initial values, and we are comparing their associated trajectories (which diverge).

Ricker Model: Chaotic Distribution

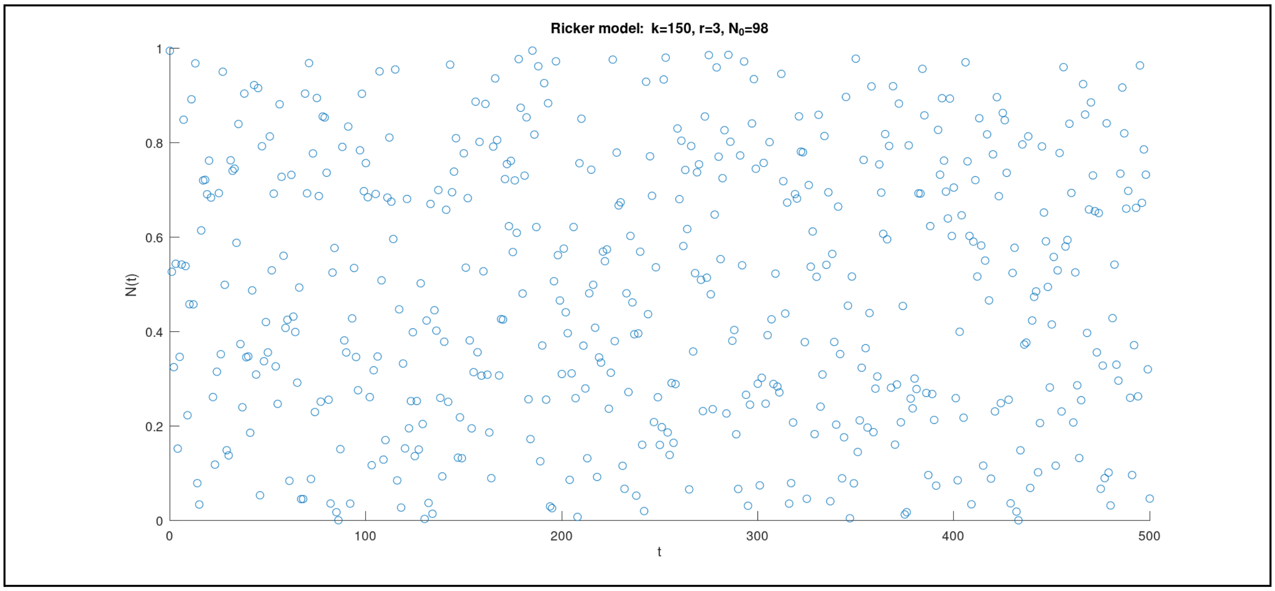


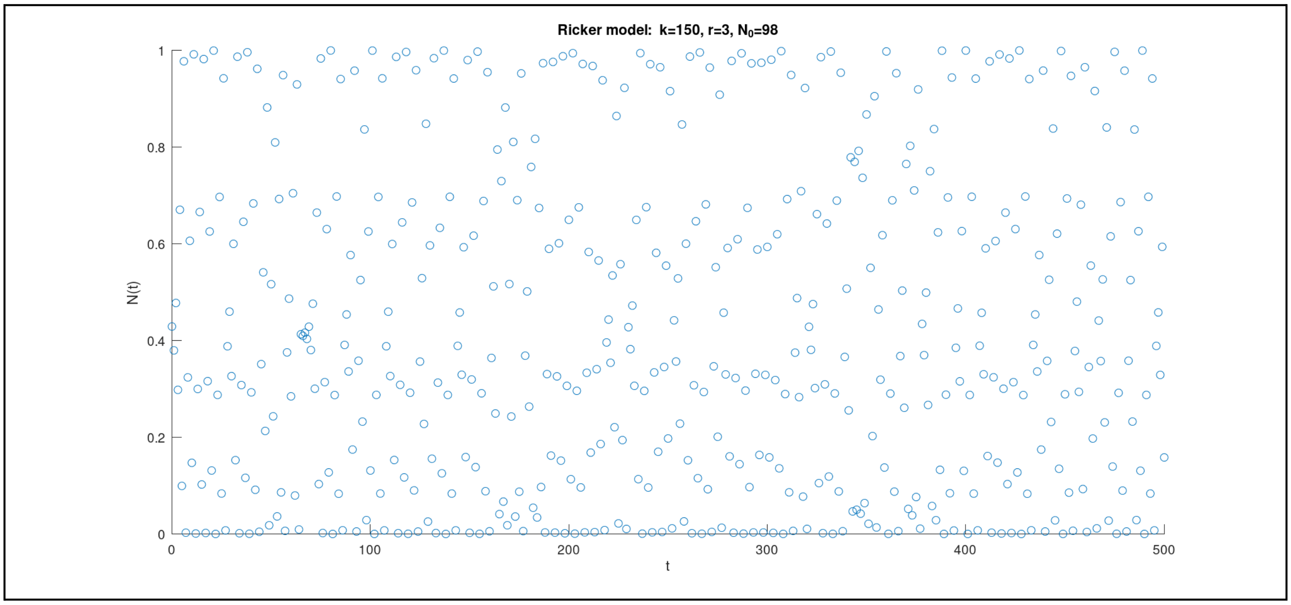
The distribution of the Ricker model appears random, but we do note some grouping towards the bottom of the graph (similar to MLM).

Ricker Model: Transformation

Using modular arithmetic we alter the Ricker Model so that its phase space exists over the unit interval. Unlike the MLM, the range of the Ricker model is not fixed – it depends on the carrying capacity k. We need to be able to vary k without changing the range, since we are using *k* as a key and our encryption scheme requires that all values lie in the unit interval. With some simple linear transformations, one can rescale and translate the distribution onto the unit interval. Alternatively, we can simply take mod(1, x_n) of all x_n. We found that this produces a more uniform (random) distribution, so we favoured this approach.

Mod1 (Top) vs Linear Transformation (Bottom)





Linear transformation:

consider vector $\mathbf{x} = (x_0, x_1, \ldots, x_n)$

length = max($\mathbf{x}$) − min($\mathbf{x}$)

c = 1 / length

$\mathbf{y} = c \cdot (\mathbf{x} - \min(\mathbf{x}))$

$\mathbf{y} = (y_0, y_1, \ldots, y_n)$ is transformation of $\mathbf{x}$ st all points lie in the unit interval

---

3 Analysis of Models

3.1- Modified Logistic Map as a Stream Key Generator

In [3] an algorithm is developed to use the MLM as a stream key generator, whereby he uses the initial condition as the stream key and iterates to create a seemingly random sequence of integers that will act as our ciphertext. This process will use a 64 bit stream as the input, the four step process is described as follows:

1. **Key Generation**

The 64 bit stream key must be converted into a value that exists on the unit interval, this is done by using the summation:

$$X_0 = \sum_{i=0}^{63} \frac{k_i}{2^{i+1}}$$

Every 64 bit key will yield its own unique value on the unit interval, which requires that the initial condition be kept secret as an eavesdropper can deduce the secret stream key from it.

2. Generation of the Lookup Table

The lookup table will be used to encrypt the letters of our plaintext, this is done by dividing the unit interval into subintervals that correspond to all the different possible characters a computer can use.

With this in mind, the unit interval is divided into 256 subintervals, each of length 1/256; each interval will represent ASCII characters. The english alphabet is represented inclusively by values between 65 and 90.

3. Encryption

The ciphertext of our encryption will be a seemingly random and divergent sequence of a finite number of iterates, which will have the same length or number of coordinates as the plaintext message represented as a vector.

Starting at the initial condition and the first letter of the message, the MLM is iterated until the value of the function lies within the interval of the corresponding letter. For example when encrypting the letter A, the function will be iterated until its value lies within the interval [65/256, 66/256) and the letter Z corresponds to the interval [90/256, 91/256).

The ciphertext value is the number of times the function is iterated until it lands in that interval, once a letter is encrypted the iteration counter is reset but the value of the function is kept the same to increase the overall security of the process.

At the suggestion of professor Frithjof we made a slight modification to the encryption method . If a letter is repeated n times, then after the first instance of the letter the ciphertext value for the remaining n-1 repeated letters will be 1. This could be a security flaw in the system, revealing where instances of repeated letters exist and giving an edge towards probability based attacks.

We changed the algorithm to skip the first counter value of 1 and continue iterating until it lands back in the same interval, thus obscuring any instance of repeated letters.

The ciphertext is the sequence of iteration values for encrypting each plaintext character.

4. Decryption

The received ciphertext can be decrypted in the same way that it was encrypted, starting with the initial condition generated in step 1. For each cipher value, iterate the MLM that number of times, then for that value of the function find the corresponding interval that lies in and find the corresponding ASCII character. This character is the

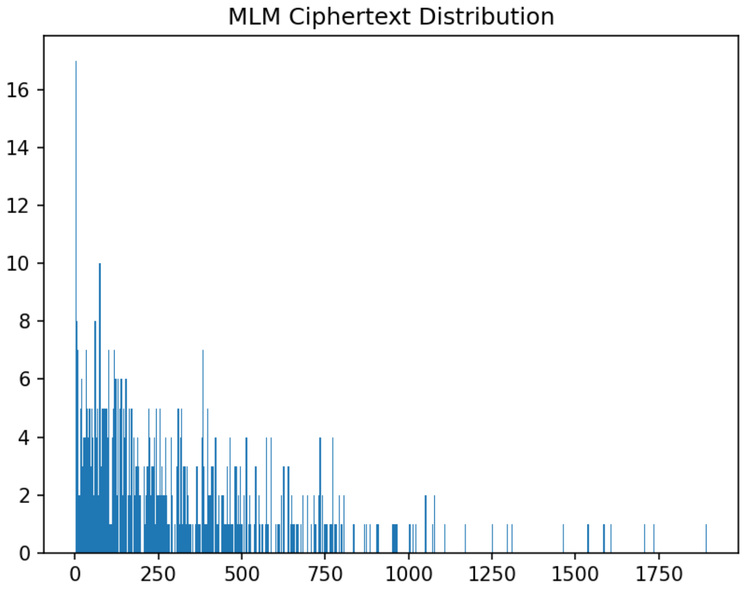
4 Contributions

4.1 Examples and Alterations of the Encryption Scheme

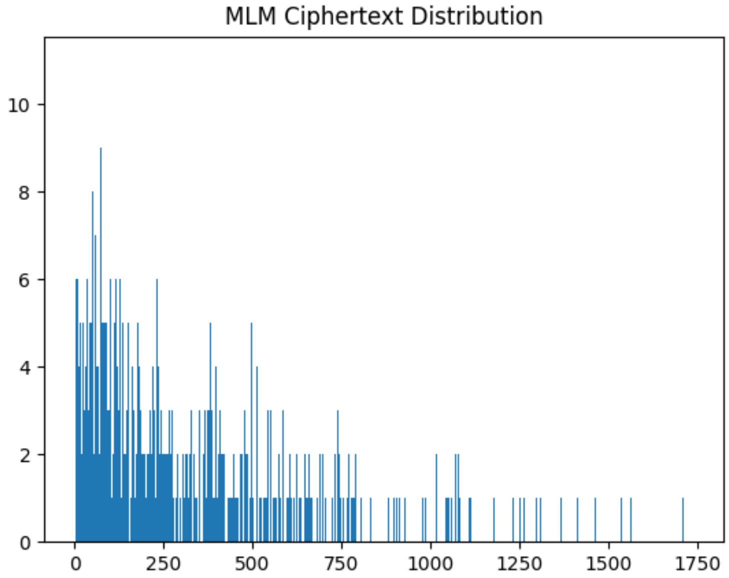
Below we demonstrate the encryption of the lyrics of Wish You Were Here by Pink Floyd, here we will show the difference between the encryption scheme outlined by Holtz and our slight variation of that which corrects the problem of repeated letters as outlined in the encryption scheme.

Example- Encrypting the Lyrics to Pink Floyd Wish You Were Here

1. The initial condition generated by our 64 bit stream key is 0.39580394849813216.
2. The lookup table is generated as it was in the description of the encryption scheme.
3. The lyrics to Wish You Were Here have 555 plaintext characters, and will yield a vector of 555 ciphertext integer values.



The distribution of ciphertext where repeated letters appear as 1 in the "random" sequence of ciphertext values. The ciphervalue 1 is repeated a total of 17 times, with *values below 200 resembling a distribution that is skewed to the left*



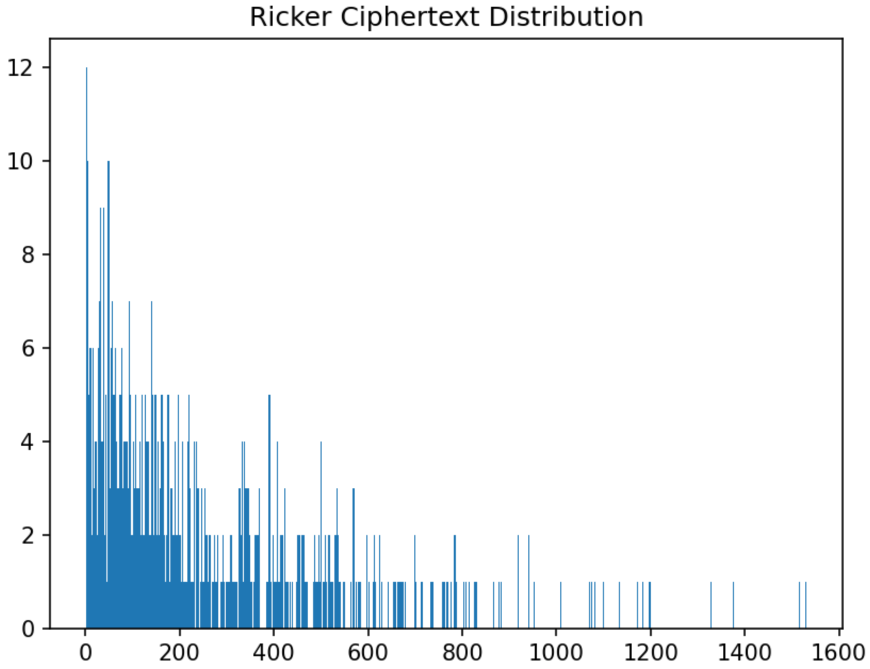
The updated version of the function, the ciphervalue of repeated letters are not represented as 1. Hence there are far fewer 1's in the distribution, with the values between 0 and 200 becoming more normally distributed.

Conclusively, our alteration of the encryption scheme shows a ciphertext distribution that reduces the efficiency of using a probability based on attack for a malicious actor to decrypt the ciphertext.

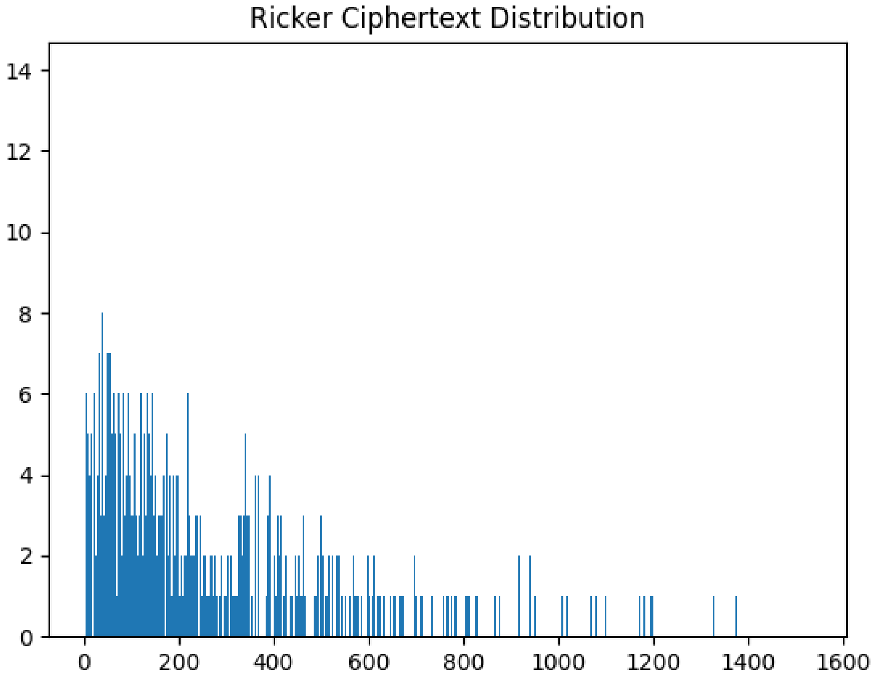
4.2 Using the Ricker Model as our Encryption Function

Our plan to extend this model is by applying this encryption scheme but using the Ricker Model in place of the MLM. As mentioned above, in section 2.3 the Ricker Model has a chaotic range for r in [2.9, 3] and k in [100,300].

Again encrypting the lyrics to Wish You Were Here by Pink Floyd and comparing the four different ciphertext distributions. That is the original encryption scheme using both the MLM and Ricker Model as encryption function, and our altered version of the encryption using both the MLM and Ricker Model.



Ricker Model ciphertext distribution of the original encryption scheme, in it the ciphertext value 1 occurs twelve times and is the most recurring value. With values less than 200 being skewed towards the left.



In our altered version of the encryption scheme the number 1 does not occur at all in the ciphertext and the ciphertext value 65 occurs the most, for a total of 8 times. It can be seen that for ciphertext values less than 200 the distribution looks similar to a uniform distribution.

5 Discussion

5.1 Results

The important differences in each ciphertext distribution can be summarized in the following table:

Model- Scheme	Probability of guessing a letter	Behaviour of Distribution for values below 200	Running time
MLM- Original	17/555	Skewed to the left	0.075 seconds
MLM- Altered	9/555	Normal Distribution	0.0608 seconds
Ricker-Original	12/555	F-Distribution, Skewed to the left	0.119 seconds
Ricker- Altered	8/555	Uniform	0.0747 seconds

Out of the four variations of the encryption scheme, the altered encryption scheme with the Ricker model as the encryption function had the best performance in this instance. It has the lowest probability of randomly guessing a word ciphertext value, the distribution for values less than 200 is the closest to being uniform, and has the fastest running time.

5.2 Limitations of Chaos on a Computer

Problems:

1. Finiteness of memory: topological spaces exist over the real numbers, because computer memory is finite it is impossible to simulate the real numbers and hence the topological property can never be truly satisfied.
2. Lack of Standardization; all data encryption methods must go through a process of standardization and validation to ensure they are viable as widely used encryption

methods.

   a. They lack standardized algorithms, key exchange protocols, and performance benchmarks.
   b. This lack of standardization makes it difficult to implement and adopt as it is required that certain standards must be met by encryption processes to ensure the safe transmission of data.

3. Performance Issues:
   a. Slow encryption speeds
   b. High computational complexity
   c. Large memory requirements, large requirements in the fact that you’re trying to simulate as much of the unit interval as you can. This is where algebraic and number theoretic ideas are superior. They work over finite fields, so there is a fixed amount of information to work with.
   d. These limitations can make it less practical to work with them in cases where efficiency and speed are necessary.
4. Lack of Awareness:
   a. Since this is a niche topic, not many cryptographers or people working in the field of data communication are aware of this form of cryptography and are reluctant to implement it since of the lack of standardization.

5.3- Conclusion

This approach to data encryption is a novel attempt to solve an important problem, ensuring that two people communicate securely. The paper and models we researched show that it is possible for a chaotic dynamic system to fulfill the Shannon requirements of secure data communication; unfortunately many issues as highlighted above plague this approach and many better alternatives within algebra, geometry, and number theory.

There is an aside I would like to highlight; this approach to encryption may not yield meaningful results or a viable system that can be used to encrypt data but it does highlight the beauty of mathematics. Two different approaches from two different fields of mathematics, algebra and dynamic systems, can solve the same problem which to us speaks as to why the topic of mathematics is so interesting.

6 References

[1] Use of chaotic dynamical systems in cryptography- Roland Schmitz

[2] R. Matthews, On the derivation of a chaotic encryption algorithm, Cryptologia XIII 1 (1989) 29–41

[3] Chaotic Cryptography: Applications of Chaos Theory to Cryptography- Nathan Holtz

[4] C.E. Shannon, Communication Theory of Secrecy Systems, Bell Systems Tech. J. 28 (1949) 656–715.

[5] Public Key Cryptography: A Dynamical Systems Perspective- Roland Schmit

[6] Edward Lorenz. Deterministic Nonperiodic Flow. March 1963. URL: https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml