

Universidad de La Habana
Facultad de Matemática y Computación



Título de la tesis

Autor: Daniel Machado Pérez

Tutores:

Dr. Roberto Mulet Genicio

Dr. Milton García Borroto

Dr. Roberto Alonso Bosch

Trabajo de Diploma
presentado en opción al título de
Licenciado en Ciencia de la Computación



La Habana,
mayo 2025

github.com/DanielMPMatCom/Thesis.git

*"La utopía está en el horizonte
Camino dos pasos y se aleja dos pasos.
Camino diez y el horizonte se corre diez pasos más allá.
Entonces, ¿para qué sirve la utopía?
Para eso, sirve para caminar."*

Fernando Birri

Agradecimientos

Resumen

Abstract

Índice general

Introducción	8
0.1. Características del <i>dataset</i> de grabaciones	8
0.2. Motivaciones de la investigación	10
0.3. Problema científico	11
0.4. Objetivos de la tesis	11
1. Detección de Individuos a Partir de Señales de Audio	14
2. Métodos y Metodologías	15
3. Resultados e Interpretación	16
Conclusiones y Recomendaciones	17
Bibliografía	18

Índice de figuras

1.	Ejemplar de <i>Eleutherodactylus eileenae</i>	9
2.	Distribución Geográfica de los Micrófonos.	10

Introducción

La rana *Eleutherodactylus eileenae* (Dunn, 1929) es una especie de la familia *Eleutherodactylidae*, endémica de Cuba. Se distribuye desde la península de Guanahacabibes, Pinar del Río, hasta la Sierra de Najasa, en la provincia de Camagüey. Ocupa hábitats como bosques semidecíduos, pinares y cafetales en altitudes de hasta 900 metros sobre el nivel del mar. Durante el día permanece oculta en oquedades de rocas y troncos, y entre la hojarasca y bromelias. Al caer la noche los machos ascienden para vocalizar en ramas y hojas de hasta 3 m de altura. [3] En la Figura 1 se muestra un ejemplar de esta especie, y en la Guía sonora de los anfibios de Cuba se puede encontrar una breve grabación de su canto y otros datos relevantes.

El canto de *E. eileenae* consta de dos notas diferenciadas: una primera “Co” breve y de frecuencia menor, seguida de “Lin” más prolongada y de frecuencia ligeramente superior, nombradas de esta forma para imitar el sonido producido. Por tal motivo y por simplicidad se procederá a nombrar “Colín” a cada uno de los cantos y también se utilizará como nombre coloquial de la especie. Estas llamadas median parámetros bioacústicos vinculados al estado fisiológico y a la competencia territorial. Además, se ha observado que los machos forman agrupaciones (“coros”), sincronizando pulsos y modulaciones para incrementar la eficacia de atracción de hembras y reducir el riesgo de depredación. La presente investigación se centrará en el estudio de dichos coros y los cantos con fines de apareamiento de los Colines.

0.1. Características del *dataset* de grabaciones

El canto de *E. eileenae* se produce en la noche, cuando los machos se agrupan para vocalizar, con picos de actividad en los meses calurosos y



Figura 1: Ejemplar de *Eleutherodactylus eileenae*.

lluviosos del año y en las horas de la tarde-noche y la madrugada. Con el objetivo de adquirir datos para los estudios de la especie, el grupo de investigación dirigido por el Dr. Roberto Alonso, de la Facultad de Biología de la Universidad de La Habana, realizó un conjunto de grabaciones de campo en la zona de la Reserva de la Biosfera “Sierra del Rosario”. Después de identificar ciertos individuos y los sitios donde usualmente se ubicaban para emitir sus cantos, se colocaron 9 micrófonos (cada uno aproximadamente a 1 metro de distancia de su Colín más cercano) cuya activación remota permitió el registro de la información acústica del coro en cuestión. Todos los dispositivos eran del mismo modelo, unidireccionales, y eran activados simultáneamente para comenzar a grabar 58 minutos consecutivos, utilizar 2 minutos para guardar los datos adquiridos y luego volver a comenzar a grabar. Esto se hizo durante 3 noches, en los períodos comprendidos entre las 18:00 horas y las 6:00 horas del día siguiente, entre los días 20 y 23 de octubre de 2023. En la Figura 2 se muestra la distribución geográfica de los micrófonos en el área de estudio. Como se puede apreciar, los micrófonos fueron colocados en un área con un radio de aproximadamente 20 metros.

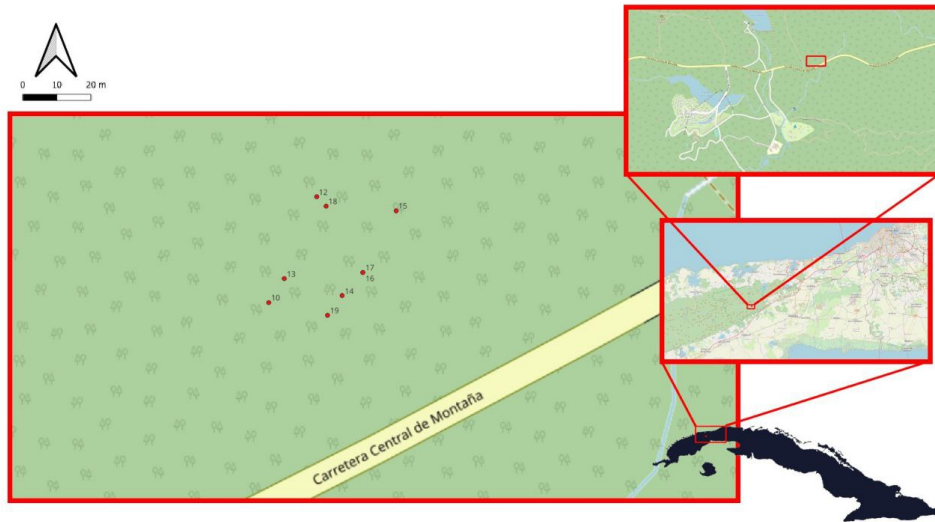


Figura 2: Distribución Geográfica de los Micrófonos.

Debido a que los micrófonos son omnidireccionales y están situados relativamente cercanos entre sí, en cada uno se registran no solamente los cantos del individuo de Colín más cercano, sino también cantos de especímenes cercanos, sonidos del ambiente natural de la zona, e incluso ruidos de agentes artificiales como autos pasando por una carretera cercana.

0.2. Motivaciones de la investigación

Como parte de los estudios sobre la ecología de esta especie de anuros se quiere profundizar en lo que se conoce sobre las características de sus cantos, de sus métodos de comunicación y su comportamiento social. Además se quiere saber si realmente se organizan en coros, y si ese fuera el caso, analizar la estructura de dichos coros, la posible existencia de un líder o protagonista, o de manera general estudiar las interacciones en el sistema formado por los Colines machos mientras cantan para atraer a las hembras.

Hasta el momento, el trabajo para procesar los audios recopilados se realizaba de forma manual con *softwares* no automáticos. La identificación

manual de cada llamado (etiquetado por individuo, hora y localización) resulta extremadamente lenta, sujeta a errores humanos y poco reproducible. Dichas razones, sumadas a la búsqueda de avances más eficientes y rigurosos, motivaron el tratamiento del problema desde un enfoque computacional, para automatizar los procesos y poder modelar matemáticamente el sistema de interacción entre los Colines.

0.3. Problema científico

El problema planteado se reduce a, dado un *dataset* de grabaciones de campo de un coro de Colines hechas con 9 micrófonos omnidireccionales, procesar los datos para obtener la información de los cantos de cada uno de los individuos implicados (momento del canto, frecuencia, energía) para analizar la estructura de dicho coro.

Que los dispositivos de sean omnidireccionales plantea la dificultad de que en una misma grabación se registra la información de sonidos que no interesan, como otros animales, autos e incluso Colines que no son los más cercanos al aparato. Por lo tanto surge la necesidad de diseñar un proceso para discriminar correctamente estos datos, y obtener una asignación Micrófono-Colín. Además es evidente que se impone encontrar una forma de eliminar el ruido y “limpiar” los audios. También se debe verificar la correcta sincronización de las grabaciones, pues a pesar de la activación remota y simultánea, el posible error de *hardware* podría comprometer la precisión de los resultados.

Convendría modelar matemáticamente el sistema para cuantificar las interacciones entre los especímenes y llevar a cabo un estudio de causalidad. Para ello se propone la utilización de un recurso clásico de la Física Estadística, el Modelo de Ising. [6]

0.4. Objetivos de la tesis

El objetivo principal de esta tesis es diseñar, implementar y validar un flujo computacional automatizado para la detección, clasificación y análisis de los cantos de apareamiento de *Eleutherodactylus eileenae*, con el fin de reconstruir la estructura de sus coros y modelar las interacciones acústi-

cas entre individuos mediante el enfoque del modelo de Ising. Para ello se plantean los siguientes objetivos específicos:

1. Construir un *dataset* limpio de grabaciones que permita extraer estadísticas fiables y facilitar la interpretación de resultados.
2. Desarrollar un método de sincronización automática de las señales grabadas por los nueve micrófonos para asegurar la consistencia temporal del *dataset*.
3. Implementar técnicas de eliminación de ruido de fondo y filtrado de eventos no deseados (otros animales, vehículos, artefactos ambientales).
4. Diseñar y comparar algoritmos para la obtención de las secuencias de cantos de cada Colín.
5. Validar la consistencia, precisión y reproducibilidad de los métodos propuestos.
6. Modelar matemáticamente la red de interacciones acústicas entre individuos utilizando el modelo de Ising:
 - Formulación del problema como Principio de Máxima Verosimilitud.
 - Implementación de algoritmos de Descenso por Gradiente para la inferencia de los parámetros J_{ij} (interacción entre el individuo i y j).
7. Analizar la idoneidad del modelo de Ising para describir la causalidad en los coros, comparando las interacciones inferidas con comportamientos observados.
8. Interpretar los resultados y extraer conclusiones ecológicas y computacionales que permitan proponer recomendaciones para estudios bioacústicos futuros.

Para abordar estos objetivos, la tesis se organiza en los siguientes capítulos:

Capítulo 1: *Detección de individuos a partir de señales de audio.* Revisión del estado del arte en detección y clasificación de llamadas bioacústicas.

Capítulo 2: *Métodos y metodologías.* Representación de audio como *mel*-espectrogramas, sincronización, eliminación de ruido, diseño de algoritmos de identificación y discriminación de Colines, y formulación del modelo de Ising con su inferencia mediante máxima verosimilitud y descenso por gradiente.

Capítulo 3: *Resultados e interpretación.* Evaluación de la sincronización automática, obtención de secuencias de cantos, hipótesis de comportamiento espectral, comparación de algoritmos, pruebas de consistencia y precisión, y análisis de las interacciones inferidas $\{J_{ij}\}$ en el modelo de Ising.

Conclusiones y Recomendaciones: Síntesis de hallazgos, discusión de implicaciones ecológicas y computacionales, limitaciones del estudio y propuestas de trabajo futuro.

Capítulo 1

Detección de Individuos a Partir de Señales de Audio

Capítulo 2

Métodos y Metodologías

Capítulo 3

Resultados e Interpretación

Conclusiones y Recomendaciones

Bibliografía

- [1] Mauricio S Akmentins. Vocal repertoire of two species of oreobates jiménez de la espada, 1872 (anura: Strabomantidae) of the yungas andean forest, nw argentina. *Journal of Natural History*, 45(29-30):1789–1799, 2011.
- [2] Mauricio Sebastián Akmentins, Laura Cecilia Pereyra, Eduardo Alfredo Sanabria, and Marcos Vaira. Patterns of daily and seasonal calling activity of a direct-developing frog of the subtropical andean forests of argentina. *Bioacoustics*, 24(2):89–99, 2015.
- [3] Roberto Alonso, Ariel Rodríguez-Gómez, and Alberto R Estrada. Patrones de actividad acústica y trófica de machos cantores de eleutherodactylus eileenae (anura: Leptodactylidae). *Revista española de herpetología*, 15(2001):45–52, 2001. (Citado en la página 8).
- [4] Daniel T Blumstein, Daniel J Mennill, Patrick Clemens, Lewis Girod, Kung Yao, Gail Patricelli, Jill L Deppe, Alan H Krakauer, Christopher Clark, Kathryn A Cortopassi, et al. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48(3):758–767, 2011.
- [5] Ryan Calsbeek, Francisco Javier Zamora-Camacho, and Laurel B Symes. Individual contributions to group chorus dynamics influence access to mating opportunities in wood frogs. *Ecology Letters*, 25(6):1401–1409, 2022.
- [6] H Chau Nguyen, Riccardo Zecchina, and Johannes Berg. Inverse statistical problems: from the inverse ising problem to data science. *Advances in Physics*, 66(3):197–261, 2017. (Citado en la página 11).

- [7] Luciano da F Costa. Comparing cross correlation-based similarities. *arXiv preprint arXiv:2111.08513*, 2021.
- [8] Almo Farina. Ecoacoustics: A quantitative approach to investigate the ecological role of environmental sounds. *Mathematics*, 7(1):21, 2018.
- [9] Hongxiao Gan, Jinglan Zhang, Michael Towsey, Anthony Truskinger, Debra Stark, Berndt J van Rensburg, Yuefeng Li, and Paul Roe. Data selection in frog chorusing recognition with acoustic indices. *Ecological Informatics*, 60:101160, 2020.
- [10] Daniel Jurafsky and James H Martin. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition.
- [11] Ignacio Sánchez-Gendríz and Linilson Rodrigues Padovese. A methodology for analyzing biological choruses from long-term passive acoustic monitoring in natural areas. *Ecological Informatics*, 41:1–10, 2017.
- [12] Jie Xie, Mingying Zhu, Kai Hu, Jinglan Zhang, Harry Hines, and Ya Guo. Frog calling activity detection using lightweight cnn with multi-view spectrogram: A case study on kroombit tinker frog. *Machine Learning with Applications*, 7:100202, 2022.
- [13] Hong-Li Zeng, Mikko Alava, Erik Aurell, John Hertz, and Yasser Rou-di. Maximum likelihood reconstruction for ising models with asynchronous updates. *Physical review letters*, 110(21):210601, 2013.
- [14] Tao Zhang, Guoqing Feng, Jinhua Liang, and Tong An. Acoustic scene classification based on mel spectrogram decomposition and model merging. *Applied Acoustics*, 182:108258, 2021.