

Sustaining Biodiversity with an Insider- Threat-Resistant Wildlife Cybercrime Investigation Tool



Timothy C. Haas, *Emeritus Associate Professor (Statistics)*

Sheldon B. Lubar College of Business, University of Wisconsin-Milwaukee, sites.uwm.edu/haas/;
Industrial Affiliate, UCLA Department of Computer Science; and
Director, Profitable Biodiversity, profitablebiodiversity.com, haas@uwm.edu

Table of Contents

- ▶ How to Sustain Biodiversity
- ▶ Biodiversity Dashboard
- ▶ A Confederation of Wildlife Crime Investigators
- ▶ Insider Threat Detection
- ▶ The New Algorithm
- ▶ Marketing a Biodiversity Offering

Biodiversity is Going Away

- The sixth mass extinction in the history of the planet is underway.
- Most large, wild mammals, many fish species, and many rare plants will be gone by 2060.
- Current conservation strategies are not working.

Sharks and Elephants

- The great white shark, a particular species of fish is endangered.
- And the African savanna elephant was added to the IUCN Red List in 2021.

And Cycads



For instance, the cycad plant, poached as a status symbol and investment, has been on this planet for about 280 million years. Dinosaurs didn't show up until 245 million years ago.

Curbing These Human-Driven Extinctions

- The wholesale killing of animals and plants needs to stop, and habitat destruction needs to be curtailed.
- Recently, wildlife crime has become the most destructive force driving species extinctions – surpassing habitat destruction for the first time.

Who is doing this killing?

- Any person engaged in the physical, acquisition of animal/plant parts for profit through the poaching (shooting, trapping, or poisoning) of live animals/plants is referred to here as a *poacher*.
- Poachers, poaching raid sponsors, and those who arrange shipments of poached animal/plant parts are called *wildlife traffickers*.

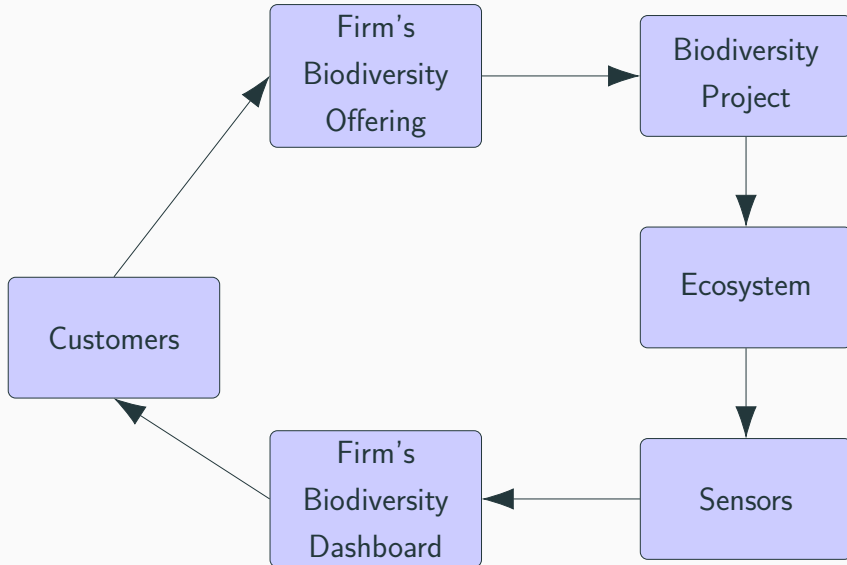
Organized Wildlife Cybercrime

- Many of these criminals organize themselves into a *wildlife trafficking syndicate* (WTS).
- A WTS conducts its business through mobile phones, social media sites, and electronic funds transfers (EFTs). Hence, wildlife trafficking is also referred to as *wildlife cybercrime*.

Private Enterprise to the Rescue

- Private, for-profit firms could save biodiversity by developing and running projects that conserve biodiversity. Indeed, private enterprise might be the last hope for many species.
- **HOW COULD THIS BE MADE TO WORK?**
In short, by making biodiversity conservation **profitable**.

Profitable Biodiversity Business Model



Selling to Biodiversity-Concerned Customers/Clients

1. A firm identifies a species they want to conserve.
2. They launch a new product or service called a **biodiversity offering** that is attached to a **biodiversity project**.
3. This project is funded by their customers/clients through a premium the firm adds to the product's sale price or to the fee charged for the service.

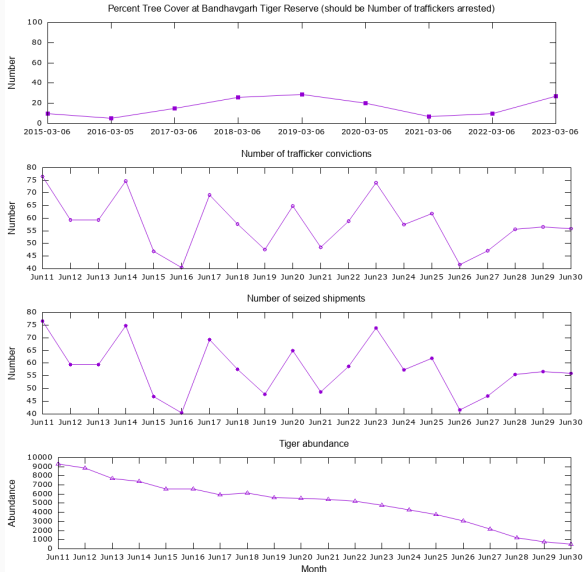
A Biodiversity Dashboard

- They earn the loyalty of these special customers/clients by maintaining a public-facing **biodiversity dashboard** that contains real-time, *audited* information on the project's progress towards conserving the targeted species.
- The JavaFX™ code that streams the needed data and creates the dashboard's graphics is at **profitablebiodiversity.com/software**.

Example: An Insurance Company

- This firm designates an existing line of auto insurance policies to be their biodiversity offering.
- Next, for their biodiversity project, they decide to participate in a *confederation* of wildlife crime investigators who are focused on tiger and tiger parts trafficking.

A Hypothetical Biodiversity Dashboard



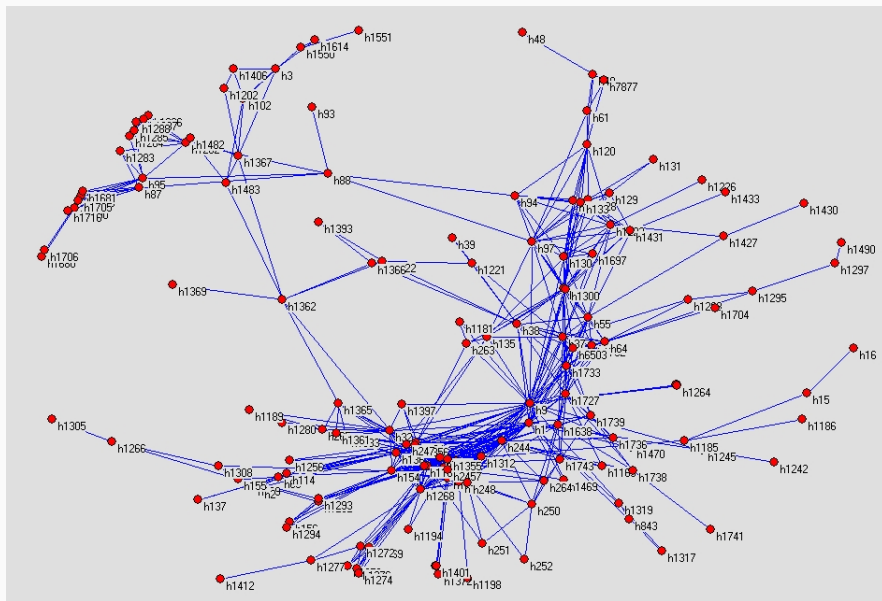
A Confederation of Trafficker Investigators and Their Database

- A confederation is proposed of wildlife crime investigators who share a *trusted* criminal intelligence database.
- This confederation analyzes intelligence from cyber and non-cyber sources to produce and then deliver to law enforcement, recommendations for detaining, surveilling, and interdicting certain traffickers.

Confederation Actions

- The confederation does this in two steps:
- **Step 1:** Compute an *actionable intelligence report*.
- **Step 2:** From this report, build the Detain, Surveil, and Interdict lists. Compute the syndicate's recovery time after those on the Detain list have been arrested.

December 2014 South African Rhino WTS



Confederation Deliverables

1. *Detain list*: A list of those WTS members (players) that law enforcement should detain for maximal disruption effect.
2. *Surveil list*: A list of those players that should be placed under surveillance for purposes of gathering evidence and/or information on future activities of the WTS.
3. *Interdict list*: A list of predicted WTS actions along with where and when these actions will take place. These actions should be interdicted.
4. *Recovery time*: An estimate of how long the WTS will take to recover from the removal of those players in the Detain list. Use this information to plan detention, surveillance, and interdiction operations.

Protecting the Confederation's Database Against Insider Attacks

- Current protocols to protect a military, terrorism, or criminal intelligence database against *insider attacks* often fail – with disastrous consequences.
- Current zero-trust cybersecurity protocols implicitly trust staff who have been granted security clearances.

Automatically Detect Insider Threats

- A modified form of the *data-centric* detection algorithm introduced by Mathew et al. (2010) is programmed into my Federated Wildlife Crime Investigations Database Management System (FWCIDBMS).
- Call this new algorithm, the *Insider Threat Detector (ITD)*.

This preprint details the ITD and gives an example.

Haas, T. C. (2026), “Sustaining Biodiversity with an Insider-Threat-Resistant Wildlife Cybercrime Investigation Tool,” Available at the Social Science Research Network (SSRN) preprint repository:
<https://ssrn.com/abstract=7238612> or
<http://dx.doi.org/10.2139/ssrn.7238612>.

Where to Download the ITD's JAVA Code

- The Zenodo repository in France:
<https://doi.org/10.5281/zenodo.19653265>
- My website (San Diego):
<https://profitablebiodiversity.com/software>

Use a Neural Network to analyze the results of a database query.

The ITD uses a *modular neural network* (MNN) classifier (Anand et al. 1995) to predict whether member j 's Structured Query Language (SQL) query results (data) are anomalous or not. The index j runs from 1 to the number of members.

Representing a Complex Query Result

- The MNN's input nodes consist of triplets of nodes, one triplet for each database attribute. The first node in a triplet is the median of the result's data on that attribute, the second node is the Interquartile Range (IQR), and the third is the number of observations on that attribute. Hence, a particular node-triplet summarizes the query result's data on its associated database attribute.
- All of these pairs form the S vector.

The MNN's Hidden and Output Nodes

- The MNN also contains at least one hidden node per output node and one output node for each confederation member.
- Output node j represents the strength of the prediction that, given the input vector's values, the query was authored by member j .

Algorithm

Each time member j sends a query to the database, the ITD algorithm runs as follows.

Step 1: Present the query result's S vector to the MNN and receive back, a prediction of the query's author. If two rules (see below) are satisfied, flag this query result as an *anomaly* for member j .

Algorithm continued

Step 2: Do this for the results from two consecutive queries sent by member j to the database.

Step 3: If both of these results are flagged as anomalous, declare member j to be an *insider threat*.

Flagging a Query Result as Anomalous

Member j 's query result is flagged as anomalous if

1. the MNN does not predict member j to be the author of the query, and
2. $(z_{norm} - z_{average})/z_{average} > p$, where z_{norm} is the value of member j 's MNN output node, and $z_{average}$ is the average of these values over all of member j 's query results in the training data set. The parameter, $p \in (0, 1)$ is a threshold. I've used $p = 0.04$.

The “Why” of These Two Rules

- The first of these rules flags a query result as being relatively different from what member j usually accesses.
- The second rule guards against false positives by requiring the query result be distinctly different from those that member j usually extracts.

First Benefit to Waiting

By waiting for two consecutive query results, the algorithm is watching for a possible temporal pattern to emerge. This aspect of the algorithm makes it particularly sensitive to *low and slow exfiltration* attacks.

Second Benefit to Waiting

In addition, waiting for two consecutive anomalous conditions to occur before declaring a member to be rogue helps to guard against a large number of *false positives*.

False positives are not good.

A false positive can unnecessarily disrupt an investigation while the confederation determines whether the member in question has actually gone rogue or not.

Just a few hidden nodes can “see” complex patterns.

The use of only two hidden nodes per output node can still cause the detector to be sensitive to a wide range of S vector patterns.

Data is king.

The result of a query rather than the query itself is the focus of this algorithm. The idea is that queries can be disguised but query results reveal what the rogue member is really after.

Marketing Analytics

Identifying the niche market:

Letting the time points be $t_1, \dots, t_J$, the model is

$$SGR_{t_j, k, l, m, n, o} = \beta_0 + t_j \beta_{t_j} + \beta_{\text{gender}_k} + \beta_{\text{age}_l} \\ + \beta_{\text{income}_m} + \beta_{\text{ownhome}_n} + \beta_{\text{education}_o} + Z_{t_j}$$

$$\text{where } Z_{t_j} = - \sum_{k=1}^p \phi_k Z_{t_j-k} + \epsilon_{t_j},$$

$$\text{and } SGR_{t_j, k, l, m, n, o} = \frac{ST_{t_j, k, l, m, n, o} - ST_{t_{j-1}, k, l, m, n, o}}{t_j - t_{j-1}}.$$

Computing Total Sales

The computed variable, $ST_{t_j, k, l, m, n, o}$ is total sales during time interval $\{t_{j-1}, t_j\}$ across all customers/clients who are in gender class k , age class l , income class m , ownhome class n , and education class o .

What drives a repeat-purchase?

Channel choice for making biodiversity offering purchase decisions through time is modeled with a multinomial, generalized logit, time series model:

$$\log \left[\frac{P(R_{i,t_i} = h)}{P(R_{i,t_i} = H)} \right] = \beta_{0,h} + t_i \beta_{t_i,h} + \beta_{\text{gender}_{i,h}} + \beta_{\text{age}_{i,h}} \\ + \beta_{\text{income}_{i,h}} + \beta_{\text{ownhome}_{i,h}} + \beta_{\text{education}_{i,h}}.$$

A SAS Code and Data Requirements

- The SAS code file, `channels.sas` fits this *autoregressive logistic regression* model to a test data set.
- Observations are needed on customers who have completed a biodiversity offering purchase. For each of these customers, observations are needed on the set of variables: $\{S_{i,t_i}, \text{channel}_{i,t_i}, \text{gender}_i, \text{age}_i, \text{income}_i, \text{ownhome}_i, \text{education}_i\}$.

A Kit and Several Codes

- The Intel Kit at **profitablebiodiversity.com** contains a set of documents that a firm can use to build their own biodiversity offering and attach to it, a biodiversity project that conducts wildlife trafficking investigations.
- All marketing analytics, database, and modeling software is at **profitablebiodiversity.com/software**

Conclusion

International wildlife trafficking is destroying biodiversity. An international confederation of wildlife crime investigators using a peer-to-peer criminal intelligence database such as the one developed herein, would help to curb this destruction by bringing these traffickers to justice.

Future Work

- Many governments and NGOs have wildlife trafficking investigation units. Demonstrate this confederation organization and associated database to them.
- Work with a private cybersecurity firm to build a confederation and associated criminal intelligence database. Acquire user impressions of its effectiveness and trust levels.