Online Appendix

1. ENUMERATION OF MODELS EVALUATED

In our paper, we evaluated the following 41 models (counting false positive cost variants, 193):

1. Positive Baseline
2. Uniform Random
3. CAPTURE
4. CAPTURE-PrevCov
5. CAPTURE-LB-PrevCov
6. CAPTURE-LB-PrevCov-DKHO
7. Standard Decision Tree
   (a) 20 False Positive Cost variants
8. BoostIT-1NN
   (a) 20 False Positive Cost variants
9. BoostIT-2NN
   (a) 20 False Positive Cost variants
10. BoostIT-3NN
    (a) 20 False Positive Cost variants
11. SVM
    (a) 20 False Positive Cost variants
12. Logistic regression
    (a) 20 False Positive Cost variants
13. Weighted Decision Trees
    (a) 20 False Positive Cost variants
14. AdaBoosted Decision Trees
    (a) 20 False Positive Cost variants
15. 3-member ensembles (18 total)
    (a) 3 voting rules × 6 combinations of decision trees
16. 5-member ensembles (9 total)
    (a) 3 voting rules × 3 combinations of decision trees

Appears in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

Classifier Type/ Test Year	F1	L&L	Precision	Recall
Decision Tree/ 2015	-	0	-	0
Decision Tree/ 2014	-	0	-	0

Table 1: Attack Prediction Results on Time Sequence Data

2. RESULTS ON TIME SEQUENCE DATA

In table 1 we provide results for the standard decision tree model when we use time-sequenced data as is used by CAPTURE and its variants. We divide the data into time period of one-year duration and therefore, one particular target can appear multiple times in the dataset with different class labels (0 or 1) for different years. We train our models on this dataset and test on both 2015 as well as 2014 test data sets. Observe that the recall and L&L scores are 0 for these learned models while precision and F1 could not be computed. This is because for all the targets in the test set, the decision tree predicts an attack with a small probability as attacks at these targets were only observed once or twice while for most of the years there were no attack observed. Therefore, the decision tree fails to predict an attack at these targets based on the thresholding of probabilities for attack/no attack. The challenge in our domain is that the data that can be collected in this domain is much more sparse compared to other crime prediction domains such as urban crime. A lot of the poaching activities cannot be recorded. Indeed, the park has a size of 2522 sq kms, and the patrollers can only cover a very small fraction of the area within a year, and an even smaller fraction of these few patrolled areas are covered more than three times. Furthermore, even while patrolling an area the rangers may not be able to observe an attack successfully due to dense habitat, for example, and hence ‘silent victims’ are unreported in the dataset. As a result, we do not have access to a detailed dataset which can enable a more comprehensive temporal analysis.

2.1 Impact of Patrol Effort on Poacher Behavior

Realizing the extent to which ranger patrols deter poachers is critical in optimizing the efficiency of wildlife protection efforts in conservation areas [1]. Understanding the underlying criminological theories [2] and exact causality relationship between rangers' coverage level and the amount of poaching activities requires controlled experiments along with a deep investigation of related environmental factors (which is not the focus of this paper). Nonetheless, our studies show that there are some promising trends in this

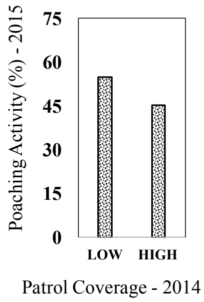

Figure 1: Poaching activity vs. previous patrol

real-world dataset: higher ranger patrol coverage amounts, in a given area, result in lower levels of poaching activities in that area.

Figure 1 demonstrates that among all of the areas that were attacked during 2015, about 55% had a LOW amount of ranger patrol coverage during the previous year (2014), and 45% had a HIGH amount of coverage. Ranger coverage is considered LOW when it is below the 50^{th} percentile and HIGH when it is in the 100^{th} percentile. These results assert that most of the areas that were subject to attack by the poachers were not sufficiently patrolled in the past. As a result, areas with LOW coverage were more attractive to poachers.

REFERENCES

[1] R. Critchlow, A. Plumptre, B. Andira, M. Nsubuga, M. Driciru, A. Rwetsiba, F. Wanyama, and C. Beale. Improving law enforcement effectiveness and efficiency in protected areas using ranger-collected monitoring data. *Conservation Letters*, 2016.

[2] M. R. Filteau. Deterring defiance:'don't give a poacher a reason to poach'. *International Journal of Rural Criminology*, 1:236–255, 2012.