Evolutionary Optimization of Epidemic Control Strategies for Livestock Disease Prevention

Krzysztof Michalak
Department of Information Technologies, Institute of Business Informatics, Wroclaw University of Economics
Wroclaw, Poland
krzysztof.michalak@ue.wroc.pl,ORCiD:0000-0003-4994-9930

ABSTRACT

This paper concerns optimizing counter-epidemic strategies selecting actions which differ with respect to the cost, damage to the animal population and the potential of stopping the disease. Important factors that have to be taken into account are the time it takes for the vaccine to become effective and the dynamic of the epidemic. This paper focuses on optimizing control strategies rather than individual decisions in the case of a particular epidemic. An evolutionary optimizer using simulations for solution evaluation is run in order to gather training data. Based on the training data a regression model is built which can subsequently be used to improve the results obtained when solving new instances of the proposed problem. Experimental results show, that the model can be used to transfer knowledge from previously solved instances to new ones.

CCS CONCEPTS

• **Mathematics of computing → Evolutionary algorithms**;

KEYWORDS

knowledge transfer; epidemics simulation models; ring vaccination

ACM Reference Format:
Krzysztof Michalak. 2019. Evolutionary Optimization of Epidemic Control Strategies for Livestock Disease Prevention. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion), July 13–17, 2019, Prague, Czech Republic.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3319619.3322069

1 INTRODUCTION

In this paper an optimization problem is studied in a scenario in which an epidemic is spreading in a livestock population. The spreading of the disease is nondeterministic with the probability of infection dependent on distances in a similar way as in the Foot-and-Mouth Disease (FMD) spreading model [1]. The goal is to optimize epidemic control *strategies*, so the optimizer adjusts parameters which are used for making decisions while the epidemic is spreading, instead of selecting a set of actions aimed at particular individuals in the population. In this paper the epidemic is simulated in a population randomly distributed on a $[0, 1] \times [0, 1]$ square using a preferential placement that results in some areas more densely and some less densely populated. In each time step, for each two individuals separated by a distance d, one infected and the other susceptible, the disease can be transmitted with the probability $P = s_0 \cdot f_{0,\sigma_T}(d)$, where s_0 is the susceptibility parameter and f_{0,σ_T} is the probability density function of a Gaussian distribution with a zero mean and a standard deviation σ_T.

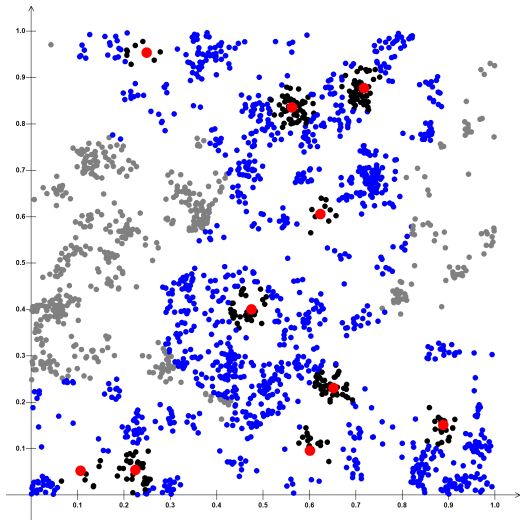

Figure 1: An example of the working of the epidemic control strategy with Θ_{cul} = 0.05 and Θ_{vac} = 0.2.

In this paper the following epidemic control strategy is defined, which uses threshold parameters Θ_{cul} and Θ_{vac}. In each time step, for each susceptible individual the minimum distance to an infected individual d_{min} is calculated. If $d_{min} <= \Theta_{cul}$ the culling action is performed and if $\Theta_{cul} < d_{min} <= \Theta_{vac}$ vaccination is performed. The effectiveness of vaccination is zero in the time step when it was administered and increases by Δ_v in subsequent time steps until the individual reaches full immunity. An example of the working of the epidemic control strategy with Θ_{cul} = 0.05 and Θ_{vac} = 0.2 is shown in Figure 1. Infected individuals are shown as large red dots, the individuals to which culling was applied as black dots and vaccinated individuals as blue dots. Gray dots represent susceptible, but as of yet, uninfected individuals.

2 PROPOSED METHOD

The method studied in this paper followed the training-testing scheme used, among others, in machine learning. For each triple of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic
© 2019 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.
ACM ISBN 978-1-4503-6748-6/19/07…$15.00
https://doi.org/10.1145/3319619.3322069

values $s_0 \in \{0.01, 0.02, \ldots, 0.05\}$, $\sigma_T \in \{0.01, 0.02, \ldots, 0.05\}$ and $\Delta_v \in \{0.1, 0.2, \ldots, 1.0\}$ (250 triples in total) the Standard Genetic Algorithm (SGA) optimized the values of $\langle \Theta_{cul}, \Theta_{vac} \rangle$ parameters. Using the optimized values a regression model $\mathcal{M}$ was trained, consisting of two polynomials in which s_0, σ_T and Δ_v were variables and one polynomial approximated Θ_{cul} and the other Θ_{vac}. For new instances of the optimization problem the regression model $\mathcal{M}$ was used to improve the solutions using two methods:

Prediction only: For the parameters s_0, σ_T and Δ_v the model $\mathcal{M}$ predicted Θ^*_{cul} and Θ^*_{vac} and no further optimization was performed. This approach is very fast, but can be expected to be inferior to the evolutionary optimizer in terms of solutions quality.

Prediction + optimization: For the parameters s_0, σ_T and Δ_v the model $\mathcal{M}$ produced predicted values Θ^*_{cul} and Θ^*_{vac} and population seeding was used in the following manner. Half of the specimens were replaced by the offspring generated by applying the Simulated Binary Crossover (SBX) [2] to a given specimen and the predicted values $\langle \Theta^*_{cul}, \Theta^*_{vac} \rangle$. Additionally, the genotype of one specimen was replaced by the predicted values $\langle \Theta^*_{cul}, \Theta^*_{vac} \rangle$.

3 EXPERIMENTS AND RESULTS

In the experiments the two methods for using information produced by the regression model presented in Section 2 were tested. For tests two sets of 30 instances with $N_p = 2000$ points were generated, one used as a training set and the other as a testing set.

To obtain training data the values of thresholds Θ_{cul} and Θ_{vac} have to be optimized for various triples of parameters s_0, σ_T and Δ_v. The difficulty in the optimization problem lies with complex simulations needed to evaluate solutions. In order to be able to perform optimization for 250 different triples of parameters (cf. Section 2) the size of the population was set to $N_{pop} = 20$ solutions and the number of generations was set to $N_{gen} = 20$. Crossover probability was set to $P_{cross} = 1.0$ and mutation probability to $P_{mut} = 1/N_p = 0.0005$. Simulated Binary Crossover (SBX) [2] was used as the crossover operator with the distribution index $\eta_c = 20$. For mutation the Polynomial Mutation [3] operator was used with the distribution index $\eta_m = 20$. Using optimized values of Θ_{cul} and Θ_{vac} as target values the regression model $\mathcal{M}$ was trained. For $\overline{\Theta_{cul}}$ a polynomial of degree 7 was obtained with the Mean Average Percentage Error (MAPE) equal 22.08%. For $\overline{\Theta_{vac}}$ a polynomial of degree 5 was obtained with the MAPE equal 18.99%.

In the testing phase the trained regression model was used to predict values of the Θ_{cul} and Θ_{vac} thresholds based on the values of s_0, σ_T and Δ_v parameters. The two approaches described in Section 2 were tested: prediction only and prediction + optimization.

Prediction only: The median cost $\overline{C^{(reg)}}$ was calculated on 30 test problem instances using a control strategy in which the regression model $\mathcal{M}$ was used to generate predicted thresholds $\langle \Theta^*_{cul}, \Theta^*_{vac} \rangle$ based on the parameters $\langle s_0, \sigma_T, \Delta_v \rangle$. For comparison, the evolutionary algorithm was run, in the exactly same manner as in the training phase, on these 30 instances and the median result $\overline{C^{(EA)}}$ was calculated.

Prediction + optimization: The median $\overline{C^{(P+EA)}}$ was calculated from costs obtained using an evolutionary algorithm in which thresholds predicted by the regression model $\mathcal{M}$ were used to augment the evolutionary optimizer as described in Section 2. For comparison, the evolutionary algorithm was run (without using predictions produced by the regression model) on these 30 instances and the median result $\overline{C^{(EA)}}$ was calculated.

In both cases the results were compared to see which median cost was lower (better) and the Wilcoxon statistical test was used to calculate the p-value of a null hypothesis stating the equality of medians. Results obtained in the tests were as follows.

Prediction only: Out of 250 medians 209 better (lower) ones were produced by the evolutionary optimizer, 40 by the regression model and 1 result was the same for both. However, in none of the cases the Wilcoxon test produced a low p-value, therefore, the difference does not seem to be statistically significant. Note, that while the evolutionary optimizer requires lots of computations, the regression model produces the prediction virtually instantaneously.

Prediction + optimization: Out of 250 medians 156 better (lower) ones were produced by the evolutionary optimizer using the regression model, 93 by the evolutionary optimizer not using the regression model and 1 result was the same for both. In the case of 31 results in favour of the evolutionary optimizer using the regression model the obtained p-value was lower than 0.05. The same was the case for only 2 results in favour of the evolutionary optimizer not using the regression model.

4 CONCLUSIONS

In this paper evolutionary optimization of epidemic control strategies was attempted with the focus on building models that could be reused for new cases of an epidemic. The proposed approach was to optimize the parameters of a control strategy using an evolutionary algorithm and use the optimized parameters as training inputs for a regression model. The trained model could subsequently be used standalone for predicting correct parameterizations or together with an evolutionary algorithm to improve its working. In this paper both approaches were tested: prediction only and prediction + optimization. The former produced results which seemed slightly inferior to those produced by the evolutionary algorithm, but in a much shorter running time. The latter approach, that is, using the regression model managed to produce competitive results in many more cases than evolutionary algorithm alone. Presented results prove that useful knowledge can be extracted from the optimization results produced by an evolutionary algorithm.

ACKNOWLEDGMENT

This work was supported by the Polish National Science Centre under grant no. 2015/19/D/HS4/02574. Calculations have been carried out using resources provided by Wroclaw Centre for Networking and Supercomputing (http://wcss.pl), grant No. 407.

REFERENCES

[1] J.A. Backer, T.J. Hagenaars, G. Nodelijk, and H.J.W. van Roermund. 2012. Vaccination against foot-and-mouth disease I: Epidemiological consequences. *Preventive Veterinary Medicine* 107, 1 (2012), 27 – 40. https://doi.org/10.1016/j.prevetmed.2012.05.012

[2] K. Deb and R. Agarwal. 1995. Simulated Binary Crossover for Continuous Search Space. *Complex Systems* 9, 2 (1995), 115–148.

[3] Kalyanmoy Deb and Mayank Goyal. 1996. A Combined Genetic Adaptive Search (GeneAS) for Engineering Design. *Computer Science and Informatics* 26 (1996), 30–45.