

## Open Problem: Online Sabotaged Shortest Path

**Wouter M. Koolen**

*Queensland University of Technology and UC Berkeley*

WOUTER.KOOLEN@QUT.EDU.AU

**Manfred K. Warmuth\***

*University of California, Santa Cruz*

MANFRED@UCSC.EDU

**Dmitry Adamskiy**

*Royal Holloway, University of London*

ADAMSKIY@CS.RHUL.AC.UK

### Abstract

There has been much work on extending the prediction with expert advice methodology to the case when experts are composed of components and there are combinatorially many such experts. One of the core examples is the Online Shortest Path problem where the components are edges and the experts are paths. In this note we revisit this online routing problem in the case where in each trial some of the edges or components are sabotaged / blocked. In the vanilla expert setting a known method can solve this extension where experts are now awake or asleep in each trial. We ask whether this technology can be upgraded efficiently to the case when at each trial every component can be awake or asleep. It is easy get to get an initial regret bound by using combinatorially many experts. However it is open whether there are efficient algorithms achieving the same regret.

**Keywords:** Online learning, combinatorial experts, specialist, sleeping.

**Introduction** Online Shortest Path is one of the core combinatorial prediction problems in online learning. In this note we consider a natural extension of this problem with sabotage. Imagine a commuter who travels from her home to her office every day. In the morning she turns on the radio and hears about all the roads that are sabotaged (due to accidents, maintenance, tornados, ...) in her area. She then chooses a route to work that avoids all sabotaged roads. Her cost is the sum of the driving times encountered along the chosen route. We assume, for simplicity that she has an app that at the end of the day obtains the driving times of all edges in the road network and can use this “full” information to improve her decisions in the future. Knowing a little bit about online learning, she proclaims her goal to be to spend almost as little time in traffic as any fixed route in hindsight. There is a snag though: how to evaluate the cost of a route that is sabotaged some of the time?

In this note we formalize this problem and call it *Online Sabotaged Shortest Path*. We then show how the use of *specialists* experts (a known tool explained below), yields a baseline answer. We argue that this baseline is rather poor: both its performance guarantee and run-time are sub-optimal. We pose improving either as intriguing open problems. We chose the iconic Online Shortest Path problem as a representative member of the family of online combinatorial decision making problems (Takimoto and Warmuth, 2003; Kalai and Vempala, 2005; Audibert et al., 2014). Any problem where the target expert is composed of components (such as matchings, fixed size subsets, spanning trees) can be extended to the case where individual components may be sabotaged.

**The Problem** Online Sabotaged Shortest Path is played on a fixed directed graph with designated source and sink nodes. Two players, learner and adversary, play a sequence of rounds. Each round the learner travels from the source to the sink, and the adversary sets the delays of all the edges.

---

\* Supported by NSF grant IIS-1118028.

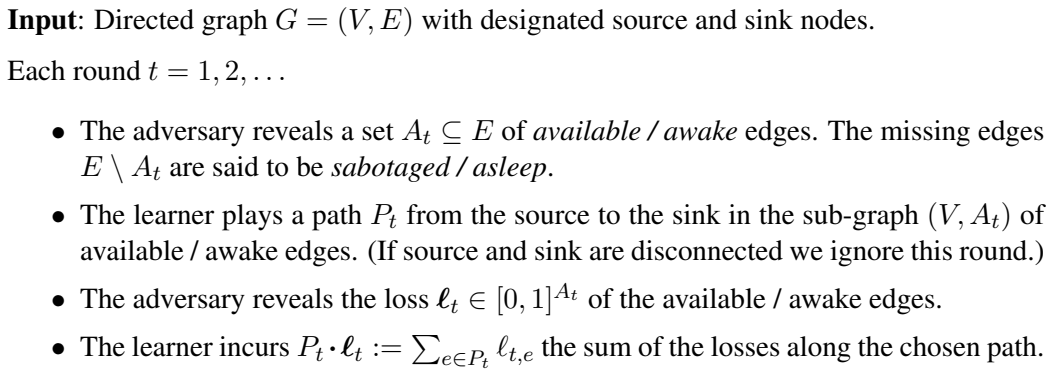


Figure 1: Online Sabotaged Shortest Path

The new ingredient is that the adversary, at the start of each round, gets to sabotage a subset of edges. More precisely, the protocol is as shown in Figure 1. As often, randomized players have the advantage, so we allow the learner to randomize and consider the resulting expected loss. The original shortest path problem is recovered when all edges are available  $A_t = E$  every round.

The objective of the learner is to incur low total loss compared to how well she could have done with hindsight. We formalize this by measuring her regret. To do so, we have to decide how to deal with sabotage. Let us say that a path  $P$  is *available / awake* if all of its edges are awake (so that a path is sabotaged if any of its edges is). We define the regret compared to a path  $P$  to be the loss of the learner minus the loss of the path  $P$ , measured on the rounds where that path was awake:

$$\text{Regret}_T(P) := \mathbb{E} \sum_{\substack{t \in \{1, \dots, T\} \text{ s.t.} \\ P \text{ awake at } t}} (P_t \cdot \ell_t - P \cdot \ell_t), \tag{1}$$

where the expectation  $\mathbb{E}$  scopes the learner’s randomized choices of the paths  $P_t$ . The goal is then to find an efficient strategy for the learner that keeps  $\text{Regret}_T(P)$  small for all reference paths  $P$  simultaneously. Note that our regret (1) differs from the ranking-based policy regret introduced by Kleinberg et al. (2010), which compares the total loss incurred by the learner to that of the best fixed policy for choosing available paths given the set of awake edges. Neu and Valko (2014) obtain efficient algorithms for policy regret in the combinatorial case under stochastic sabotage, but Kanade and Steinke (2014) prove hardness under adversarial sabotage. We are looking for intuitive but less expressive objectives like (1) that render the problem tractable in full adversariality.

**Initial Approach Using Specialists** In this section we approach Online Sabotaged Shortest Path using the *specialist* framework introduced by Freund et al. (1997). Specialists are an extension of prediction with expert advice where the experts (now called specialists) are allowed to abstain (or sleep). One application of specialists is where experts are themselves algorithms that may not complete within a fixed time budget. We first briefly review the basic regret bound from specialists, and then reduce sabotage to specialists to get a first (horribly!) rough algorithm and regret bound.

The specialists problem is an extension of the vanilla  $N$ -experts problem. To deal with specialists the canonical multiplicative weight algorithm is generalized slightly: the awake specialist weights are updated multiplicatively and renormalized within the awake experts, the weights of asleep experts are not changed. This scheme guarantees (see Freund et al. 1997; Cesa-bianchi et al.

2012)<sup>1</sup> that the the regret compared to each reference specialist  $i \in \{1, \dots, N\}$  is bounded by  $\mathbb{E} \sum_{t:i \in A_t} (\ell_{t,i_t} - \ell_{t,i}) \leq \sqrt{T \ln N}$  where the sum is taken over the rounds where expert  $i$  was awake, and the expectation  $\mathbb{E}$  scopes the algorithm’s random choice of expert  $i_t$ . Note that this regret has the same shape as (1). The algorithm runs in time  $O(N)$  per round and uses  $O(N)$  space.

Sabotage can be reduced to specialists quite naturally as follows. We enumerate all source-sink paths  $P_1, \dots, P_D$  (note that  $D$  can be exponential in the size of the graph). We then use these paths as specialists, where a path  $P$  is awake at time  $t$  iff all its edges  $e \in P$  are awake  $e \in A_t$ . Disregarding gross inefficiency, the specialists algorithm would deliver the following regret bound. For any comparator path  $P$ , the expected regret (1) is bounded by  $\text{Regret}_T(P) \leq K\sqrt{T \ln D}$  where  $K$  is the length of the longest path (i.e. the loss range).

**Discussion** This regret bound has two issues. First, we know that its regret bound is sub-optimal. In the simple case where no edges are sabotaged, the Component Hedge algorithm (instance of mirror descent) improves the regret by a factor  $\sqrt{K}$  for many graphs. This is due to the so-called *range factor problem* discussed by Koolen et al. (2010) that arises from the fact that the algorithm does not take advantage of the fact that paths may overlap, and that their losses cannot be controlled independently. Second, and perhaps more importantly, the running time is horrible as one parameter is maintained per path. We have efficient algorithms for combinatorial settings without sleeping. These algorithms are “collapsed”: they maintain one parameter per edge. The question is hence whether and how this methodology can be extended to incorporate sabotaged / sleeping edges.

We believe that Online Sabotaged Shortest Path a core problem which will become a starting point for various orthogonal extension such as switching between paths, semi-bandit feedback, etc.

## References

- Jean-Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. Regret in online combinatorial optimization. *Math. Oper. Res.*, 39(1):31–45, 2014.
- Nicolò Cesa-bianchi, Pierre Gaillard, Gabor Lugosi, and Gilles Stoltz. Mirror descent meets fixed share (and feels no regret). In *Advances in Neural Information Processing Systems 25*, pages 980–988, 2012.
- Yoav Freund, Robert E. Schapire, Yoram Singer, and Manfred K. Warmuth. Using and combining predictors that specialize. In *Proc. 29th Annual ACM Symposium on Theory of Computing*, pages 334–343, 1997.
- Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.
- Varun Kanade and Thomas Steinke. Learning hurdles for sleeping experts. *Transactions on Computation Theory*, 6(3):11:1–11:16, 2014.
- Robert Kleinberg, Alexandru Niculescu-Mizil, and Yogeshwar Sharma. Regret bounds for sleeping experts and bandits. *Machine Learning*, 80(2-3):245–272, 2010.
- Wouter M. Koolen, Manfred K. Warmuth, and Jyrki Kivinen. Hedging structured concepts. In *Proceedings of the 23rd Annual Conference on Learning Theory*, pages 93–105, 2010.
- Gergely Neu and Michal Valko. Online combinatorial optimization with stochastic decision sets and adversarial losses. In *Advances in Neural Information Processing Systems 27*, pages 2780–2788, 2014.
- Eiji Takimoto and Manfred K. Warmuth. Path kernels and multiplicative updates. *Journal of Machine Learning Research*, 4:773–818, 2003.

---

1. Fancier bounds with mixture comparators and relative entropy are also available, but outside the scope of this note.