Attack Design for Continuous and Discrete Systems

DJ Passey

BYU Information and Decision Algorithms Laboratories
Professor: Sean Warnick
March 20th, 2018

Introduction

Feedback loops in dynamic systems give rise to a certain type of vulnerability. Self-regulating systems and controlled systems are often susceptible to attack. When an attacker has access to the signals passing through a link and also has the ability to adjust the signal, the attacker may have the ability to destabilize the entire system with small, undetectable, signal adjustment.

Example of Vulnerability: Power Grids

In the summer of 1996, California, Nevada, Utah and several other western states experienced region wide blackouts. For a time it was unclear why. Suprisingly, analysts traced the outages to a routine failure of a single unimportant line. They were perplexed. It was almost unthinkable that a problem so small could spiral into 7.5 million people without power. [**?**]

Because of these blackouts, the Western Systems coordinating council discovered that their power network concealed a hidden vulnerability. Though the network seemed robust enough to handle ordinary failures, an ordinary failure in just the right spot crippled the power grid. Why did this happen? The failures occured because of the structure of the power grid and because of system's self regulation—because of it's feedback structure.

The computerized controler regulating the power grid responed to failures by rerouting current to nearby lines. It was in feedback with the power grid, taking measurements and making adjustments. Unfortunately, because of the structure of the network, when the controler responded to the failure of the vulnerable line, it rerouted the power to another overtaxed line. This line failed and double the current was rerouted to a third line that also failed. Soon enough, the surge was too strong for any line or group of nearby lines to handle and it swept through the western states resulting in the collapse of the entire system.

It is important to note that the computerized controller is normally very effective at handling outages and failures. In ordinary conditions it is able to keep the power grid stable. However, there was a structural weakness in the network, and when the controller tried to handle the failure in a key part of the network, it led to the destabilization of most of the power grid.

Many systems have weaknesses in their network structure and controler strategies, weaknesses that may not be visible to the arbiters of such systems. Using the tools of control theory, we can discover these hidden vulnerabilities and shore them up or use them in a targeted attack.

What makes systems vulnerable?

In order to design an attack on a feedback system, it is necessary to understand why feedback systems have vulnerabilities and how to isolate them. We begin with a linear time invariant model of a system

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

and solve for its transfer function $G(s) = C(sI - A)^{-1}B + D)$ such that:

$$y = G(s)u$$

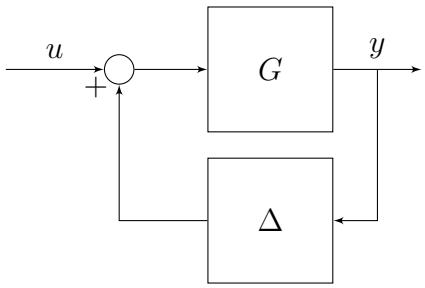

Figure 1: An attaker Δ in feedback with a stable system G

Assume that G is stable. Let Δ represent an attacker that is in feedback with G. (Figure 1). We interpret this as Δ reading in the output of G and adding on an adjustment to the inputs of G accordingly.

We can imagine an input u signal passing through G and generating an output $y = Gu$. This output is read in to Δ and $\Delta\, y = \Delta Gu$ is added on to the original signal u. The loop then repeats with the input $u + \Delta Gu$ passing through G, then through Δ with the result being added on to u. The loop continues indefinitely and as such we represent the input to G as the series

$$u + \Delta Gu + \Delta G\Delta Gu + \cdots$$

Producing

$$y = G\ [I + (\Delta G) + (\Delta G)^2 + (\Delta G)^3 \cdots]u$$

$$y = G\ [\sum_{k=0}^{\infty} (\Delta G)^k]u$$

By the small gain theorem, the series $\sum_{k=0}^{\infty}(\Delta G)^k = (I - \Delta G)^{-1}$ and converges if $||\Delta G|| < 1$. [**?**] Then the sum diverges when

$$||\Delta G|| \geq 1 \text{ or } ||\Delta|| \geq \tfrac{1}{||G||}$$

Thus, the size of the smallest attack capable of destabilizing G is,

$$||\Delta|| = \frac{1}{||G||}$$

We see here, that because of the structure of a feedback loop, small problems are repeatedly incoorperated into the input and output. These small problems can grow into to large unbounded issues. Taking advantage of this property, we can design a minimal attack Δ capable of destabilizing a system. We are guarenteed that such a Δ exists. [?] Therefore, it is always possible to construct a minimal attack on the system G.