Cyber Attack in a Two-Area Power System: Impact Identification using Reachability

Peyman Mohajerin Esfahani, Maria Vrakopoulou, Kostas Margellos,
John Lygeros and Göran Andersson

***Abstract*— This paper presents new results on the applications of reachability methods and computational tools to a two-area power system in the case of a cyber attack. In the VIKING research project a novel concept to assess the vulnerabilities introduced by the interaction between the IT infrastructure and power systems is proposed. Here we develop a new framework and define a systematic methodology, based on reachability, for identifying the impact that an intrusion might have in the Automatic Generation Control loop, which regulates the frequency and the power exchange between the controlled areas. The numerical results reveal the weaknesses of the system and indicate possible policies that an attacker could use to disturb it.**

I. INTRODUCTION

The electric power trasmission system is probably the most vital infrastructure in our society [1]. Large power systems are nowadays very complex and tightly coupled with the SCADA system, which supervises them in terms of collecting data from remote facilities and sending back control commands. The resilience of power system on this infrastructure, makes it more susceptible not only to operational errors but also to external attacks.

The SCADA system measures data through remote devices and transmits them to control centers through communication channels. There computer processing takes place and control commands are sent back to the system. The vulnerabilities that are introduced could be exploited by malicious attackers. In [2], [3], [4], [5] real examples of cyber attacks were reported. The authors of [6] proposed a framework in order to clarify the interaction between the power system and the IT infrastructure and identify the vulnerabilities and the malfunctions of both that could lead to an abnormal operation of the power network. From another perspective the authors of [7] attempted to quantify the impact of a cyber attack in the power market.

The work in this paper is motivated by the framework proposed by the VIKING research project [8]. This project proposes a novel concept to address the challenges introduced by the interaction between the IT systems and the power transmission and distribution systems. Main objective is to identify the vulnerabilities of these safety critical infrastructures, determine the impact that possible failures or attacks might have and develop strategies to mitigate these effects.

Here we investigate the impact of a cyber attack on the Automatic Generation Control (AGC) in a power system. The primary objective of the AGC is to regulate frequency to the specified nominal value and maintain the power exchanged between the controlled areas to the scheduled values by adjusting the generated power of specific generators in the area. AGC actions are usually determined for each control area at a central dispatch center. Measured system frequency and tie line flows are sent to this center and then a feedback signal that regulates the generated power is sent back to the generators, participating in the AGC, through the SCADA system.

AGC is one of the few control loops that are closed over the SCADA system without human operator intervention. To reveal its vulnerabilities, we consider a two-area power system and analyze the safety of the system in the case where an attacker has gained access to the AGC signal of one area and is able to inject any undesirable input to it. For this purpose, a dynamic nonlinear frequency model, which is suitable for load-frequency studies for the two interconnected areas, was developed.

To determine whether there exists a signal for the attacker that could irreversibly disturb the system, we perform an analysis based on reachability methods. Reachability for continuous and hybrid systems has been an important topic of research in the dynamics and control literature. A wide range of applications from air traffic management systems [9], [10], [11] and flight control [12], [13] to biology and economics have been formulated in the framework of reachability theory. To the best of our knowledge though, only a few problems in power systems have been studied in the context of reachability [14], [15], [16]. In [14] the underlying system was hybrid with both differential and algebraic equations and the authors used reachability tools in order to investigate a voltage safety problem. In the same context [16], fault release control of a double machine-infinite bus system was studied.

One common way of addressing reachability questions is by using optimal control methods. In this case, the value function of an appropriate optimal control problem, is the viscosity solution to a first order partial differential equation in the standard Hamilton-Jacobi form [17], [18], [19]. Reachable sets can then be computed, using tools like [20], [21], [22] based on level set methods [23], [24].

Peyman Mohajerin Esfahani, Kostas Margellos and John Lygeros are with the Automatic Control Laboratory, Department of Electrical Engineering, Swiss Federal Institute of Technology (ETH), Physikstrasse 3, ETL I22, 8092, Zürich, Switzerland. email: {mohajerin, margellos, lygeros}@control.ee.ethz.ch

Maria Vrakopoulou and Göran Andersson are with the Power Systems Laboratory, Department of Electrical Engineering, Swiss Federal Institute of Technology (ETH), Physikstrasse 3, ETL G26, 8092, Zürich, Switzerland. email: {vrakopoulou, andersson}@eeh.ee.ethz.ch

This paper has two main contributions. It develops a framework in order to quantify the safety of a two-area power system, but also provides a new methodology so as to identify, and in future work evaluate, the impact that a cyber attack could have. The latter is based on a reachability formulation, and enables us to construct the policy that an attacker could follow so as to disturb the system.

In Section II the physical description and the mathematical model of the two-area power system is presented. Section III provides the formulation of the problem in the reachability framework. In Section IV simulation results that validate our approach are presented. Finally in Section V we provide some concluding remarks and direction for future work.

II. PHYSICAL DESCRIPTION AND MATHEMATICAL MODELING

A. Modeling of the Two-Area Power System

Consider the system of Fig. 1, which consists of two interconnected control areas, each one equipped with its own AGC, connected by a tie line of reactance X. Following [25], [26], each area is approximated by an equivalent generating unit G_i equipped with primary frequency control.

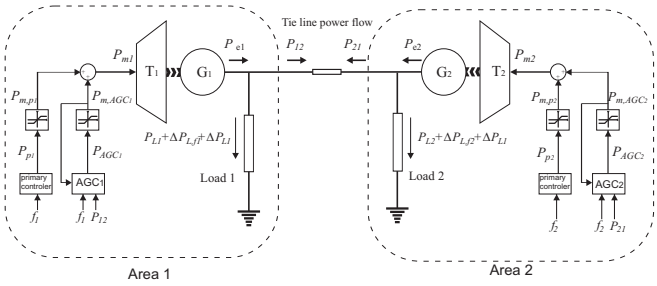

Fig. 1. Two-Area Power System with AGC

In each area, in stable steady-state conditions, the total generated active power, which is assumed to be equal to the mechanical one (P_{m_i}), is the same as the power demand (i.e area loads and exchanged power). Any disturbance of this balance results in a deviation of the system frequency from its set-point value. The behavior of the frequency in the area i with respect to the power balance can be approximated by

$$\Delta \dot{f}_i = \frac{f_0}{2H_i S_{B_i}} (\Delta P_{m_i} - \Delta P_{e_i}), \tag{1}$$

where f_i is the frequency of the area i (Hz), H_i denotes its inertia time constant(sec MW/MVA), P_{m_i}, is the generated power (MW), P_{e_i} is the consumed power (MW), and S_{B_i} is the power base (MVA). The 0 index stands for the nominal value of each variable and the Δ operator denotes the deviation from its nominal value.

After an increase in the power demand P_{e_i}, the rotating parts of the generators will start losing their kinetic energy until the point that the consumed and the produced power are equal and a new equilibrium is reached due to the frequency dependency of the load. However, this stabilizing effect is normally too small to be able to keep the frequency within reasonable bounds. Therefore, to keep the frequency deviation at an acceptable level, generators are equipped with a regulating unit (governor) that performs automatic primary frequency control. The primary frequency control law is given by $\Delta P_{pi} = -\frac{1}{S_i}\Delta f_i$, where the proportional gain S_i is referred to as speed droop or speed regulation. Since a proportional controller is used for this task, after the activation of the governor the frequency does not return to its nominal value. Also, in an interconnected system with two or more independently controlled areas, the scheduled power interchange between these areas will not be respected after the response of the primary control.

Supplementary control action is needed to ensure that the frequency and power exchange return to their nominal values; this is provided by the AGC. The AGC of the area i is typically a proportional-integral (PI) controller. To avoid wind up in case of saturation, an anti-wind up circuit is also used [27]. The overall block diagram for the AGC of a single area is shown in Fig. 2, where C_{p_i}, T_{N_i} are parameters of the AGC model of area i and K_{a_i} is a constant of the anti-wind up circuit.

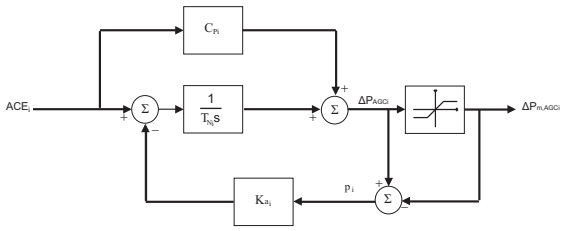

Fig. 2. PI controller with anti-wind up

The Area Control Error signal (ACE_i), is defined as $ACE_i = \Delta P_{ij} + B_i \Delta f_i$ where P_{ij} is the power transmitted from area i to area j, $P_{0_{ij}}$ denotes its scheduled value and $B_i = \frac{1}{S_i}$ according to Non-Interactive Control [25].

A change in the consumed power P_{e_i} in area i is expressed through $\Delta P_{e_i} = \Delta P_{L_i} + \frac{1}{D_{l_i}}\Delta f_i + \Delta P_{ij}$, where ΔP_{L_i} is the actual deviation of the load, the second term refers to the deviation due to the frequency dependence of the load and the last one due to power exchange between the two areas.

It should be also noted that the power flow on the tie-line from area 1 to area 2 is described by $\Delta P_{12} = P_T \sin(\Delta\phi)$, where it is assumed that the scheduled transferred power in the tie line is 0 ($P_{0_{12}} = 0$). Note that the active power losses on the line are neglected so $\Delta P_{21} = -\Delta P_{12}$ and P_T is the maximum transfer power on the line which is assumed to be constant (i.e. the grid topology is not changed). $\Delta\phi$ is the voltage angle difference between the ends of the line and is related to the frequencies through $\Delta\dot{\phi} = 2\pi(\Delta f_1 - \Delta f_2)$.

Consider now the case of a cyber attack in the second area. We assume that the attacker has disabled the AGC in this area and is using its input U_d i.e. $\Delta P_{AGC_2} = U_d$ to introduce a bounded disturbance to our system. According to the previous analysis and considering the saturations introduced in Fig. 1, the model of the two-area power system for this case could be described by the following set of differential equations

$$
\begin{aligned}
\Delta\dot{f}_1 &= \frac{f_0}{2H_1 S_{B_1}}\Big(\Delta P_{m,p_1} + \Delta P_{m,AGC_1} - \frac{1}{D_{l_1}}\Delta f_1 - P_T \sin\Delta\phi\Big),\\
\Delta\dot{f}_2 &= \frac{f_0}{2H_2 S_{B_2}}\Big(\Delta P_{m,p_2} + \mathbf{U}_d - \frac{1}{D_{l_2}}\Delta f_2 + P_T \sin\Delta\phi\Big),\\
\Delta\dot{\phi} &= 2\pi(\Delta f_1 - \Delta f_2),\\
\Delta\dot{P}_{AGC_1} &= \Big(\frac{1}{D_{l_1}}\frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}} - \frac{1}{S_1}\frac{1}{T_{N_1}}\Big)\Delta f_1\\
&\quad - \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}}\Delta P_{m,p_1} - \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}}\Delta P_{m,AGC_1}\\
&\quad - \Big(\frac{1}{T_{N_1}} - \frac{C_{p_1} f_0}{2S_1 H_1 S_{B_1}}\Big)P_T \sin\Delta\phi\\
&\quad - 2\pi C_{p_1} P_T(\Delta f_1 - \Delta f_2)\cos\Delta\phi - \frac{K_{a_1}}{T_{N_1}} p_1. \qquad (2)
\end{aligned}
$$

where due to saturation we have

$$
\Delta P_{m,p_i} = \begin{cases} \Delta P_{p_i}^{min} & \text{if} \quad \Delta P_{p_i} \le \Delta P_{p_i}^{min} \\ \Delta P_{p_i} & \text{if} \quad \Delta P_{p_i}^{min} < \Delta P_{p_i} < \Delta P_{p_i}^{max} \\ \Delta P_{p_i}^{max} & \text{if} \quad \Delta P_{p_i} \ge \Delta P_{p_i}^{max} \end{cases}
$$

$$
\Delta P_{m,AGC_1} = \begin{cases} \Delta P_{AGC_1}^{min} & \text{if} \quad \Delta P_{AGC_1} \le \Delta P_{AGC_1}^{min} \\ \Delta P_{AGC_1} & \text{if} \quad \Delta P_{AGC_1}^{min} < \Delta P_{AGC_1} < \Delta P_{AGC_1}^{max} \\ \Delta P_{AGC_1}^{max} & \text{if} \quad \Delta P_{AGC_1} \ge \Delta P_{AGC_1}^{max} \end{cases}
$$

$$
p_1 = \begin{cases} 0 \ \text{if} \quad \Delta P_{AGC_1}^{min} < \Delta P_{AGC_1} < \Delta P_{AGC_1}^{max} \\ \Delta P_{AGC_1} - \Delta P_{m,AGC_1} \qquad \text{else} \end{cases}
$$

S_{B_i}	f_0	H_i	D_{l_i}	S_i	C_{p_i}	T_{N_i}
$10\ GW$	$50Hz$	$50\ s$	$\frac{1}{200} MW/Hz$	$0.002 Hz/MW$	0.1	30
$\Delta P_{AGC_i}^{max}$	$\Delta P_{AGC_i}^{min}$	$\Delta P_{p_i}^{max}$	$\Delta P_{p_i}^{min}$	P_T	K_a	
$350MW$	$-350MW$	$75\ MW$	$-75\ MW$	$1000\ MW$	100	

TABLE 1: Parameters of area i of our system.

Since we are interested in the impact of unreasonable changes of the AGC signal, no changes at the actual load of the areas is considered ($\Delta P_{L_i} = 0$) in the above model. In our computations, we assumed the two areas to be equal and hence used the same data for both of them. In the reachability analysis that follows the state vector x is defined as $x = [x_1\ x_2\ x_3\ x_4]^T = [\Delta f_1\ \Delta f_2\ \Delta\phi\ \Delta P_{AGC_1}]^T$.

B. Safety Considerations

The frequency control outlined in the previous subsection is vital to the satisfactory performance of the power system. The controllers try to keep the frequency to its nominal value because too large deviations could damage the power system devices. This action may in the end jeopardize the stability of the whole system and in the worst case lead to a system blackout. In normal operation Δf should not exceed 1.5Hz.

The amount of power that a line can transfer is also limited to maintain reliability and stability in the system. The limiting value for the permissible power transfer is influenced, according to the line length, by three factors: the thermal limit, the voltage drop and the stability limits.

In the case-study of the two-area system, the amount of power that can be transferred is considered to be limited only by the steady state stability limit. This limit is a percentage of the maximal power P_T. We consider a minimum allowable steady state margin as 30% [26] which implies that $\Delta P_{12} \in [-70\%P_T, +70\%P_T]$. Since P_T is assumed constant, a bound $\Delta\phi \in [-44°, 44°]$ in the phase difference is considered.

In summary we consider the system to be safe when the state trajectories of (2) lie inside the following safe set of the state space:

$$
\begin{aligned}
\Delta f_1 &\in [-1.5, +1.5]\\
\Delta f_2 &\in [-1.5, +1.5]\\
\Delta\phi &\in [-44°, 44°]
\end{aligned} \qquad (3)
$$

Moreover it should be noted that large power oscillations in the tie-line are undesirable and can lead to triggering out-of-step relays that trip generating units in order to avoid potential damaging mechanical vibrations [26]. This would be another type of disturbance that a cyber attacker could try to excite by his intrusion.

III. REACHABILITY FORMULATION

The model of the two-area power system described in the previous section is a continuous, nonlinear control system of the form $\dot{x} = f(x, u)$, with $x \in \mathbb{R}^n$, $u \in U \subseteq \mathbb{R}^m$, and $f(\cdot,\cdot) : \mathbb{R}^n \times U \to \mathbb{R}^n$. Let $\mathcal{U}_{[t,t']}$ denote the sets of Lebesgue measurable functions from the interval $[t, t']$ to U. Following the notation of (17), if U is compact, f is Lipschitz in x and continuous in u, and $T \ge 0$ is an arbitrary time horizon, then this system with initial condition $x(t) = x \in \mathbb{R}^n$ admits a unique solution $x(\cdot) : [t, T] \to \mathbb{R}^n$ for all $t \in [0, T]$, $x \in \mathbb{R}^n$, $u(\cdot) \in \mathcal{U}_{[t,T]}$. For $\tau \in [t, T]$ we will use $\phi(\tau, t, x, u(\cdot)) = x(\tau)$ to denote this solution.

Main objective would be to characterize the set of initial states for which there exists at least one signal for the attacker that at some time within the time horizon, can lead the system trajectory outside the safe set K. This is a typical reachability problem [17], and can be solved by computing the set

$$
\begin{aligned}
Inv(t, K) = \{x \in \mathbb{R}^n \mid \forall u(\cdot) \in \mathcal{U}_{[t,T]}\\
\forall \tau \in [t, T]\ \phi(\tau, t, x, u(\cdot)) \in K\}.
\end{aligned}
$$

If $x \in Inv(0, K)$ then whatever the attacker does, the system will not leave the safe set over the horizon $[0, T]$. Otherwise, if $x \notin Inv(0, K)$, there exists a control policy for the attacker to steer the system outside the safe set at some time within the internal $[0, T]$. The computation of the set $Inv(t, K)$ can be formulated as an optimal control problem. To eliminate technical difficulties, we assume that K is closed and is given as the zero level set $K = \{x \in \mathbb{R}^n \mid l(x) \ge 0\}$ of the bounded, uniform continuous function $l(\cdot) : \mathbb{R}^n \to \mathbb{R}$. A reasonable choice for the function l is the signed distance to the set K, i.e. $l(x) = -d(x, K)$ if $x \in K^c$ (K^c denotes the complement of K), and $l(x) = d(x, K)$ if $x \in K$. $d(x, K) = \inf_{\hat{x}\in K} \|x - \hat{x}\|$ stands for the usual distance to a set. We can then introduce the value function $V : \mathbb{R}^n \times [0, T] \to \mathbb{R}$

$$
V(x, t) = \inf_{u(\cdot)\in\mathcal{U}_{[t,T]}} \min_{\tau\in[t,T]} l(\phi(\tau, t, x, u(\cdot))), \qquad (4)
$$

where the objective of the input $u(\cdot)$ is to minimize the minimum value attained by the function l along the state

trajectory over the horizon $[t, T]$. It can be shown [17], [18] that $Inv(t, K) = \{x \in \mathbb{R}^n \mid V(x, t) \geq 0\}$.

Moreover, V is the unique, bounded and uniformly continuous viscosity solution to the Hamilton-Jacobi equation

$$\frac{\partial V}{\partial t}(x,t) + \min\{0, \inf_{u \in U} \frac{\partial V}{\partial x}(x,t) f(x,u)\} = 0, \tag{5}$$

with terminal condition $V(x, T) = l(x)$.

A second reachability concept that we will use for the analysis of the AGC system is that of viability [28]. Viability characterizes the set of initial states for which there exists at least one input for the attacker that can keep the system trajectory in a given set K within the specified time horizon. This set could be defined as

$$Viab(t, K) = \{x \in \mathbb{R}^n \mid \exists u(\cdot) \in \mathcal{U}_{[t,T]} \\ \forall \tau \in [t, T]\ \phi(\tau, t, x, u(\cdot)) \in K\}.$$

Similar to the invariance analysis, the viability computation could also be formulated as an optimal control problem and to eliminate technical difficulties, we assume that K is open and is related to a function l by $K = \{x \in \mathbb{R}^n \mid l(x) > 0\}$.

We can then introduce the value function $\bar{V} : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$

$$\bar{V}(x,t) = \sup_{u(\cdot) \in \mathcal{U}_{[t,T]}} \min_{\tau \in [t,T]} l(\phi(\tau, t, x, u(\cdot))), \tag{6}$$

where the objective of the input $u(\cdot)$ is to maximize the minimum value attained by the function l along the state trajectory over the horizon $[t, T]$. It can be shown [17], [18] that $Viab(t, K) = \{x \in \mathbb{R}^n \mid \bar{V}(x, t) \geq 0\}$, where $\bar{V}$ is the unique, bounded and uniformly continuous viscosity solution to the Hamilton-Jacobi equation.

$$\frac{\partial \bar{V}}{\partial t}(x,t) + \min\{0, \sup_{u \in U} \frac{\partial \bar{V}}{\partial x}(x,t) f(x,u)\} = 0, \tag{7}$$

with terminal condition $\bar{V}(x, T) = l(x)$.

IV. SIMULATION AND RESULTS

In this section, we adopt the reachability framework of Section III to provide answers to some questions concerning the safety of the two-area power system. For this purpose, it is examined if the attacker, by applying a suitable control policy, is able to violate the safety constraints (3). Considering different bounds on the attack signal, it will be shown that for sufficiently large attack bounds, the power system is vulnerable to such cyber attacks. The case of violating the power exchange constraint between the two areas is investigated and the analysis proves the existence of an attack strategy that might directly or indirectly lead to a swinging in the exchange power. These scenarios have been tested numerically by the Level Set Method Toolbox of [22].

A. Power Exchange Range Violation

In this part, we will examine if the attacker could construct a policy so as to exceed the $\Delta\phi$ bounds. Therefore, we define K_1 as

$$K_1 := \{x \in \mathbb{R}^4 \mid |x_1| \leq 1.5, |x_2| \leq 1.5, |x_3| \leq 44^\circ\}, \tag{8}$$

and the distance function $l(x) = \min\{x_1 + 1.5, 1.5 - x_1, x_2 + 1.5, 1.5 - x_2, x_3 + 44^\circ, 44^\circ - x_3\}$. Note that the last state x_4, which corresponds to the AGC signal in the first area is restricted indirectly due to the line saturation.

For this safety analysis, we performed a series of reachability computations for different bounds of the attack input. Fig. 3 shows a family of curves that correspond to the different bounds of the attack signal. These curves quantify how the volume of the safe set changes in time. By obtaining this figure, we can conclude that the attacker would need a signal at least $(200MW)$ to disturb the system starting from the nominal operating point.

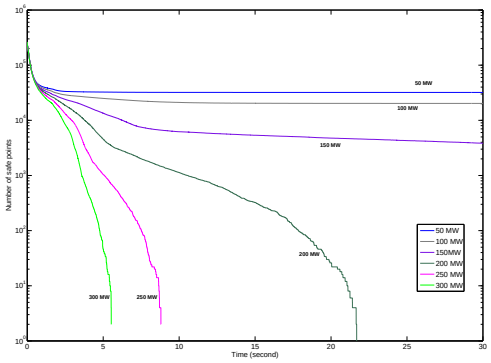

Fig. 3. Volume of the safe set for different bounds of the attack signal

In the sequel, we depict the case that the attacker is able to inject an arbitrary signal up to the 100 MW; i.e. $|U_d| \leq 100$MW. In Fig. 4 as also expected from Fig. 3, it is clear that the safe set has saturated after approximately 10 seconds, which means that despite the attack, there are still some states, including the nominal point, that system trajectories can start and remain in the safe region K_1 of (8).

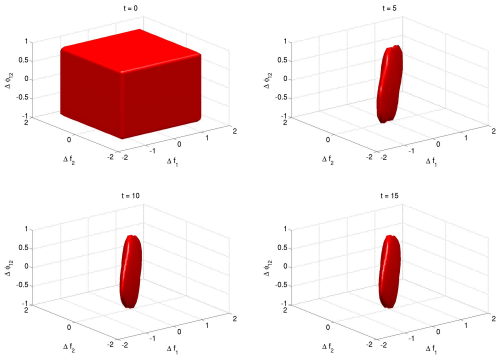

Fig. 4. Safe set for the case where $U_d \in [-100, +100]MW$ and $\Delta\phi \in [-44^\circ, 44^\circ]$

B. Power Swinging Between Two Areas

The reachability analysis that we performed guarantees the existence of a control policy, which at some time could

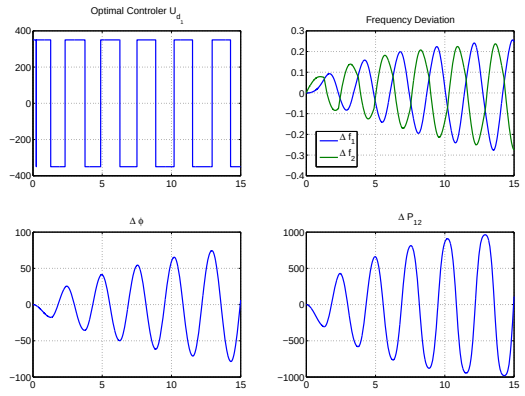

Fig. 5. (*a*) Example of a control policy for the case where $U_d \in [-350, +350]MW$ and $\Delta\phi \in [-44°, 44°]$; (*b*) frequency deviation in the two areas;(*c*) angle ϕ deviation; (*d*) power in the tie-line

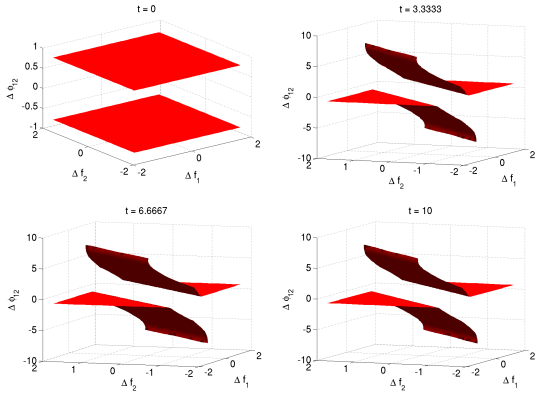

Fig. 6. Viability computation

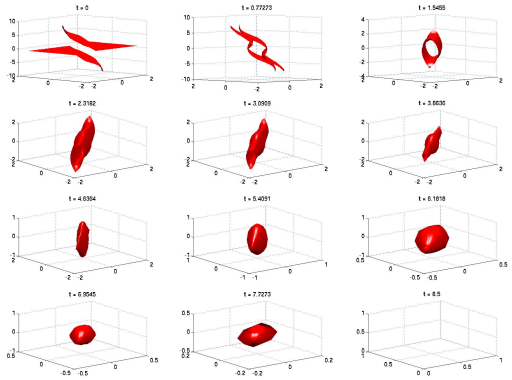

Fig. 7. Reachability computation

lead the system trajectories outside K_1. We can observe that the angle deviation between the two areas is the state that violates the constraints first. But there is no guaranty that the attacker could find a signal so as to keep $\Delta\phi$ above 44° for a sufficiently large amount of time. Nevertheless, since the invariant set is empty (for $U_d \geq 200MW$), the attacker will always be able to construct a sequence of policies to repeatedly force the system to exceed 44°. In Fig. 5, an input signal that leads to such behavior is illustrated. The swinging in the angle $\Delta\phi$ results in a similar performance of the power exchange between the two areas. This swinging, could activate the so called out-of-step protection relays, which in turn would trip the generators. Note that the ΔP_{12} oscillations in Fig. 5.*a* correspond to a frequency almost equal to the resonance frequency of the two-area system which is given by

$$f_r = \frac{1}{2\pi}\sqrt{\frac{\pi f_o P_T(H_1 + H_2)}{S_B H_1 H_2}} \simeq 0.4Hz, \tag{9}$$

where $S_B = S_{B_1} = S_{B_2}$ since the areas were considered to be equal.

Next we consider the possibility of keeping the angle $\Delta\phi$ above 44° for a sufficiently large amount of time. We define a new set K_2 according to the angle constraints as

$$K_2 := \{x \in \mathbb{R}^4 \mid |x_3| > 44°\}. \tag{10}$$

As explained earlier, the solution to this problem follows from (7) with K_2 as the terminal set. In Fig. 6, it was assumed that the attacker signal is bounded in [−350 350] MW due to the AGC saturation. One could see that the viability set is saturated in approximately 7 seconds; namely, there exists a non-empty set such that if the system starts from that set, the attacker could construct an input sequence so as to keep the angle over 44° (or below −44°) for a specified time horizon. Notice that since the other states (except $\Delta\phi$) are free in this case, the constraint ($|x_3| > 44°$) in the definition of K_2 divides the state space to two parts; one part between the two surfaces of Fig. 6, and one outside. The latter is the set where the attacker is trying to steer the system trajectories.

To determine whether the attacker could keep the angle deviation within certain bounds, or to increase it unboundedly over time, we restrict the set K_2 defined in (10) as follows

$$K_3 := \{x \in \mathbb{R}^4 \mid 180° > |x_3| > 44°\},$$

and repeat the same viability analysis, which now shows that the set gets empty in less than 2.5 seconds. In other words, the angle deviation will increase indefinitely, since there is no way to keep it bounded in K_3.

It remains to verify whether the attacker is able to force the system to that set. If so, then once reaching the viability set, the attacker could change his control policy and keep the angle deviation in the unsafe region for sufficiently large amount of time. For this purpose, we define the set K_4

$$K_4 := \{x \in \mathbb{R}^4 \mid \bar{V}(x, 0) > 0\}, \tag{11}$$

where $\bar{V}(x, 0)$ is the value function obtained from the viability computation.

As shown in Fig.7, after an invariant calculation, for every initial condition, there exists at least one control policy for the attacker so as to reach the viability set in 8.5 seconds. Then, the attacker could switch policy and keep the state trajectory in K_3.

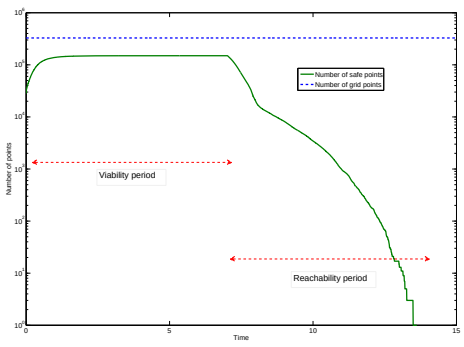

Fig. 8. The dash line is the total number of grid points; the solid line presents the volume of the safe set over time

Fig. 8 summarizes the previous analysis, which comprises of two stages; the first stage provides a way to compute the viability set K_3, whereas the second describes the computation of K_4. One can see how the volume of the safe part of the state space changes. At the 7th second the viability set is saturated, and the attacker could change policy so as to keep the angle increasing. That way, based on analysis of section II, the power will start swinging and this in turn might lead to the out of step relay triggering.

V. CONCLUSION

A reachability framework to perform safety analysis for a two-area power system was developed. This analysis proposed a new methodology so as to identify how an attacker could disturb the system by gaining access to the AGC, and determine the policy that he should follow to disrupt the system. Our approach was tested numerically, by using computational tools based on reachability.

In this paper, we assumed that the attacker has access to all system states whereas in reality it might not be the case. A nonlinear observer, so that the attacker is able to estimate the states that cannot directly measure, has been already developed and will be used in future work in a more realistic set-up. We will also work on fault detection schemes to determine whether it is possible to diagnose the attacker's action sufficiently fast before he disturbs the system.

VI. ACKNOWLEDGMENT

The authors would like to thank Prof. D. Kirschen for helpful discussions. This research work is supported by the European Commission under the project VIKING, FP7-ICT-SEC-2007-1.

REFERENCES

[1] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 Major Grid Blackouts in North America and Europe, and Recommended Means to Improve System Dynamic Performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.

[2] *Forbes, Congress Alarmed at Cyber-Vulnerability of Power Grid, available at http://www.forbes.com/2008/05/22/cyberwar-breach-government-tech-security_cx_ag_0521cyber.html.*

[3] *CNN, Sources: Staged cyber attack reveals vulnerability in power grid, available at http://www.cnn.com/2007/US/09/26/power.at.risk/index.html.*

[4] *Comptuterworld, DHS to review report on vulnerability in West Coast power grid, available at http://www.computerworld.com/s/article/9138017.*

[5] J.-W. Wang and L.-L. Ronga, "Cascade-based attack vulnerability on the US power grid," *Elsevier, Safety science*, vol. 47, no. 10, pp. 1332–1336, 2009.

[6] D. Kirschen and F. Bouffard, "Keep the Lights On and the Information Flowing," *Power and Energy Magazine, IEEE*, vol. 7, no. 1, pp. 50–60.

[7] M. Negrete-Pincetic, F. Yoshida, and G. Gross, "Towards Quantifying the Impacts of Cyber Attacks in the Competitive Electricity Market Environment," *IEEE Power Tech Conference*, 2009.

[8] *VIKING Project, http://www.vikingproject.eu.*

[9] C. Livadas, J. Lygeros, and N. Lynch, "High level modelling and analysis of the trafic alert and collision avoidance system (tcas)," *Proceedings of the IEEE*, vol. 88, no. 7, pp. 926–948, 2000.

[10] C. Tomlin, J. Lygeros, and S. Sastry, "A game theoretic approach to controller design for hybrid systems," *Proceedings of the IEEE*, vol. 88, no. 7, pp. 949–969, 2000.

[11] C. Tomlin, I. Mitchell, and R. Ghosh, "Safety verification of conflict resolution manoeveures," *IEEE Transactions on Intelligent Transportation Systems*, vol. 2, no. 2, pp. 110–120, 2001.

[12] J. Lygeros, C. Tomlin, and S. Sastry, "Controllers for reachability specifications for hybrid systems," *Automatica*, vol. 35, pp. 349–370, 1999.

[13] M. Oishi, C. Tomlin, V. Gopal, and D. Godbole, "Addressing multi-objective control: Saftety and performance through constrained optimization," *In M.Di.Benedetto and A.Sangiovanni-Vincentelli, editors, Hybrid Systems: Computation and Control*, pp. 459–472, 2001.

[14] I. M. Mitchell and Y. Susuki, "Level Set Methods for computing Reachable Sets of Hybrid Systems with Differential Algebraic Equation Dynamics," *Lecture Notes in Computer Science (LNCS)*.

[15] J. Licheng, L. Haifeng, R. Kumar, V. Ajjarapu, J. McCalley, N. Elia, and V. Vittal, "An Application of Reachable Set Analysis in Power System Transient Stability Assessment," *Power Engineering Society General Meeting, IEEE*, 2005.

[16] Y. Susuki and T. Hikihara, "Predicting Voltage Instability of Power System via Hybrid System Reachability Analysis," *in Proceedings of the American Control Conference*, pp. 4166–4171, 2007.

[17] J. Lygeros, "On reachability and minimum cost optimal control," *Automatica*, vol. 40, no. 6, pp. 917–927, 1999.

[18] L. Evans and P. Souganidis, "Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations," *Indiana University of Mathematics Journal*, vol. 33, no. 5, pp. 773–797, 1984.

[19] I. Mitchell, A. M. Bayen, and C. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE transactions on Automatic Control*, vol. 50.

[20] ——, "Validating a Hamilton Jacobi approximation to hybrid reachable sets," *in M.Di.Benedetto and A.Sangiovanni-Vincentelli (eds.) Hybrid Systems: Computation and Control Springer Verlag*, pp. 418–432, 2001.

[21] I. Mitchell and C. Tomlin, "Level set methods for computations in hybrid systems," *in M.Di.Benedetto and A.Sangiovanni-Vincentelli (eds.) Hybrid Systems: Computation and Control Springer Verlag*, pp. 310–323, 2000.

[22] I. Mitchell, "Application of level set methods to control and reachability problems in continuous and hybrid systems," *Stanford University, PhD thesis*, 2002.

[23] J. A. Sethian, *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. New York: Cambridge University Press, 1996.

[24] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[25] G. Andersson, *Dynamics and Control of Electric Power Systems*. ETH Zürich, 2009.

[26] P. Kundur, *Power System Stability and Control*. McGraw-Hill Inc., 1994.

[27] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*. Prentice Hall, 2002.

[28] J. Aubin, *Viability Theory*. Boston:Birkhauser, 1991.