This is the accepted version of a paper presented at The 4th IEEE International Conference on Smart Grid Communications (SmartGridComm) in Vancouver, Canada, 21-24 October, 2013.

Citation for the original published paper:

Vuković, O., Dán, G., Bobba, R. (2013)
Confidentiality-preserving Obfuscation for Cloud-based Power System Contingency Analysis.
In: Proceedings of IEEE SmartGridComm, October 2013

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-134821
Confidentiality-preserving Obfuscation for Cloud-based Power System Contingency Analysis

Ognjen Vuković György Dán
KTH Royal Institute of Technology
Stockholm, Sweden
Email: {vukovic,gyuri}@kth.se

Rakesh B. Bobba
University of Illinois at Urbana-Champaign
Urbana, IL, USA.
Email: rbobba@illinois.edu

Abstract—Power system operators are looking to adopt and migrate to cloud technologies and third-party cloud services for customer facing and enterprise IT applications. Security and reliability are major barriers for adopting cloud technologies and services for power system operational applications. In this work we focus on the use of cloud computing for Contingency Analysis and propose an approach to obfuscate information regarding power flows and the presence of a contingency violation while allowing the operator to analyze contingencies with the needed accuracy in the cloud. Our empirical evaluation shows, i) that the errors introduced into power flows due to the obfuscation approach are small, and ii) that the RMS errors introduced grow linearly with the magnitude of obfuscation.

I. INTRODUCTION

Power grids around the world are undergoing a transformation to accommodate more renewable generation, allow consumer interaction with the infrastructure, and improve efficiencies through modernization. At the heart of this transformation are new sensor deployments, such as smart meters and phasor measurement. These new sensors are producing large volumes of data that a power system operator has to process and store, and increasing the number of devices that a power utility has to connect to and manage. To manage the data and connectivity to these devices utilities are looking to cloud based services. Smart meters in particular, given their numbers (in Millions even for medium size Utility) and geographic distribution, pose a challenge. Responding to this demand, many companies (e.g., GE’s GRID IQ, Honeywell’s Akuacom, AutoGrid etc.) are offering cloud-based software-as-a-service models to manage smart meters and associated applications such as automated Demand Response (DR). Apart from customer facing applications such as Demand Response, utilities are also looking into leveraging cloud computing for other services such as managing security of their infrastructure as evidenced by the new CIGRE working group (D2.37) on cloud technologies for managed security [1]. The primary drives towards cloud computing are lower costs, improved efficiencies and elasticity of computing provided.

Power system applications related to operations, such as Contingency Analysis, forecasting, Optimal Power Flow, etc., could also benefit from the advantages cloud technologies provide [2]). Security and reliability concerns are, however, a major barrier for adopting cloud technologies for power system operations [2], [3], especially with third-party providers. Recent work has addressed this issue from two sides. First, by improving the reliability and security provided by the cloud infrastructure for power grid applications, as is being done in the GridCloud [4], [5] project. Second, by transforming power system applications to preserve security properties such as confidentiality, integrity and availability in third party infrastructures. Borden et al., [6] focus on transforming the optimal power flow problem before instantiating it in the cloud to preserve confidentiality.

In this paper we focus on contingency analysis (CA), which is a core application in power system operation. A contingency corresponds to the failure of one or more system components, such as a transmission line, a transformer, or a generator. The failure of any of these components would lead to a change in the power flows on the transmission lines, and could potentially result in an unstable system (e.g., power flows that exceed the thermal capacity of transmission lines). The aim of contingency analysis is to determine whether the power system would be unstable in case any of a potentially large set of contingencies would happen.

Contingency analysis is performed in modern energy management systems every time a new state estimate becomes available as a result of state estimation - as often as every few minutes. The number of contingencies that needs to be considered depends on the instantaneous load of the power system, the higher the load the more contingencies might have to be considered. The number of contingencies considered in practice is limited by the computational power available in the control center, and is often constrained to considering the loss of single components known as $N-1$ security. Cloud-based contingency analysis could allow an operator to scale the number of considered contingencies freely as a function of the instantaneous system state and enable $N-\varepsilon$ security that is considered desirable, but it could expose the current system state and possible critical contingencies, thereby facilitating targeted attacks.

In this paper we propose an algorithm to obfuscate information regarding power flows and the presence of a contingency violation while allowing the operator to analyze contingencies with the needed accuracy in the cloud. We show that our approach doesn’t introduce any error for CA using DC model. Further our empirical evaluation shows that the error introduced by the approach when using an AC model is quite small and that RMS error grows linearly with the magnitude of obfuscation applied.

The rest of the paper is organized as follows. Section II provides necessary background on contingency analysis. Section III presents our adversary model and usage scenario and
Section IV describes our obfuscation approach. Section V discusses some preliminary evaluation results and Section VI concludes the paper.

II. BACKGROUND

We consider a power system that consists of $N$ buses. We denote by $P_n$, $1 \leq n \leq N$ the power injection (load or generation) at bus $n$, and $P_I$ is the vector of power injections. We denote the state of the power system by $x$. For simplicity, we consider active power flows only, in which case the system state is determined by the phase angles at the buses, and thus $x$ is the vector of phase angles.

Given the system state $x$, the power flow between buses $n$ and $m$ can be computed as

$$P_{nm} = V_n V_m (G_{nm} \cos x_{nm} + B_{nm} \sin x_{nm}) = f_{nm}(x_{nm}),$$

where $x_{nm} = x_n - x_m$ is the phase angle difference between buses $n$ and $m$, and $G_{nm}$ and $B_{nm}$ are the real and imaginary parts of the bus admittance matrix corresponding to buses $n$ and $m$. The power injections can be computed using Kirchhoff’s nodal law, and we denote the power injections as a function of the system state by $P_I = f_I(x)$. Finally, one can express the vector of power injections and power flows as a function of the system state as $P = f(x)$

A. AC Load-flow based Contingency Analysis

Let $c$ be a contingency (e.g., the failure of two transmission lines), and let $f^c$ be the function that describes the power flows under contingency $c$ as a function of the system state, i.e., $P^c = f^c(x)$. Observe that a contingency might change the system topology and thus $f^c(\cdot) \neq f(\cdot)$. Similarly, the vector of power injections $P^c_I$ under contingency $c$ might be different from $P_I$, e.g., if the contingency involves the loss of one or more generators. To describe the relationship between the power injections before and after the contingency we introduce the fault matrix $F^c$ such that $P^c_I = F^c P_I$. If contingency $c$ does not affect the power injections then $F^c$ is the identity matrix.

Given the vector of power injections $P^c_I$ under contingency $c$, contingency analysis requires the solution of the load-flow problem, i.e., finding the state vector $x^c$ that solves $P^c_I = f^c(x^c)$. The state vector is obtained through solving the power balance equations,

$$\Delta P_n = -P_n + \sum_{m} P_{nm} = 0. \tag{2}$$

Since the sum of the injections over all buses is zero, there are in total $N - 1$ power balance equations and $N - 1$ unknowns, as the phase angle of the reference bus is set to zero.

The equations (1) are non-linear, thus the solution to (2) is obtained using an iterative numerical method, typically the Newton-Raphson method [7]. Starting from an initial guess $x^c(0)$, the Newton-Raphson method obtains an updated estimate at iteration $k$ by computing

$$\Delta x^c(k + 1) = -J_k^{-1}\Delta P_I(k), \tag{3}$$

where $J_k = \frac{\partial P_I}{\partial x} |_{x=x^c(k)}$ is the Jacobian evaluated at the most recent guess $x^c(k)$, and then letting $x^c(k+1) = x^c(k) + \Delta x^c(k+1)$. Observe that the Jacobian is a non-singular square matrix of size $(N - 1 \times (N - 1)$. The algorithm terminates when the power mismatch $\Delta P_I$ is below a certain threshold. Let $x^c$ be computed system state under contingency $c$.

Given the system state $x^c$ under the contingency, the power flows can be calculated as $P^c = f^c(x^c)$. If any of the power flows exceeds the capacity limit (e.g., thermal capacity) of the transmission line then the system is said to be in an insecure state, and a corrective action must be taken by the operator to move the system to a state in which no contingency results in a capacity violation.

III. ADVERSARY MODEL AND SCENARIO

A. Adversary Model

We assume that the adversary has knowledge about the topology of the system but that he doesn’t have access to the current state of the system. That is he does not know what the instantaneous power injections and power flows are. This adversarial model is inline with the recent body of work on false data injection attacks (e.g., [8]-[10]) where the adversary is assumed to have full or partial knowledge of the H matrix for a DC model.

The goal of the adversary is to find the current system state (flows and injections) so he can determine if there are any contingencies with critical violations. Correspondingly, the goal of the obfuscation algorithm is to mask the real power flows from the adversary and to hide the existence of a violating contingency.

B. Usage Scenario

As shown in Figure 1, when a power system operator wants to undertake CA he will create an obfuscation vector and send the system with obfuscated flows to the cloud for contingency analysis. On obtaining the result of the CA for the various contingencies, the operator performs a deobfuscation step to obtain the power flows and injections that correspond to the non-obfuscated (actual) system. While obfuscation is performed only once, deobfuscation is performed for every contingency. Nevertheless, much of the computation of the deobfuscation can be done a-priori for a particular system topology.

IV. OBFUCSTED CONTINGENCY ANALYSIS

In the following we first introduce the proposed obfuscation algorithm. We then show that for DC load flow calculation the proposed obfuscation does not introduce an error.

A. Obfuscation Algorithm

Consider the known power injections $P^c_I$ under a contingency $c$. If an adversary has access to the power injections $P^c_I$ and the computed power flows $P^c$ under the contingency, it can infer which part of the system is most critical for stability and could perform a targeted attack. It is therefore important to obfuscate the information exposed to a potential attacker.

In the following we propose an algorithm that limits the attacker’s ability to infer potential system instability. We do so by obfuscating the system state on which contingency analysis
is performed, and by compensating the contingent system for the modification after contingency analysis is performed. The important property of the proposed algorithm is that the computational cost of the obfuscation and of the de-obfuscation is much less than that of the contingency analysis.

1) Obfuscation: Given $P$, the actual power flows in the system, obfuscation consists of adding a randomly chosen vector of power flows to the actual power flows. We refer to the latter as the \textit{power flow obfuscation} vector,

$$P^o = Hx^o, \quad (4)$$

where $H = \frac{\partial P}{\partial x}$ is the Jacobian evaluated at the most recent system state (prior to the CA), and $x^o$ is a non-negative vector of phase angles, the \textit{state obfuscation} vector. We will discuss in Section IV-C how to create the state obfuscation vector. We use the state obfuscation vector to create the obfuscated system state, $x^O = x + x^o$. The obfuscated system state can be used to compute the obfuscated power injections as

$$P^O = f(x^O). \quad (5)$$

The obfuscated system state, and the corresponding obfuscated power injections $P^O_i = f_i(x^O)$ are the basis for the contingency analysis performed in the cloud.

For a particular contingency $c$, the obfuscated power injections $P^O_i$ are created, and are used as the input to the non-linear load-flow problem. The solution to the load-flow problem, i.e., the result of the analysis for contingency $c$ is the state $x^{c,O}$ of the obfuscated contingent system.

2) Deobfuscation for a Contingency: Given the result $P^{c,O} = f_o(x^{c,O})$ of the contingency analysis performed on the obfuscated power flows for contingency $c$, deobfuscation consists of compensating for the power flows introduced through obfuscation.

To describe deobfuscation we define $H_c = \frac{\partial P^c}{\partial x}$, the Jacobian of the system under contingency $c$ evaluated at the most recent system state (as in (4)). The deobfuscated power flows under contingency $c$ are then obtained as

$$\tilde{P}^c = P^{c,O} - H_cJ_c^{-1}P^{c,O}_i, \quad (6)$$

where $P^{c,O}_i$ is the vector of obfuscation power injections under contingency $c$. Note that if the contingency involves the loss of a generator then at least one or two entries in $P^{c,O}_i$ are changed and thus $P^{c,O}_i \neq P^i$. Due to the non-linearity of the power balance equations, obfuscation will introduce an error in the result of the contingency analysis. We quantify this error by the difference of the power flows under a contingency with and without obfuscation

$$e_P = P^c - \tilde{P}^c. \quad (7)$$

To express the relative error we furthermore define the maximum componentwise relative error

$$e_P = \max_i \frac{e_{P,i}}{P^i}, \quad (8)$$

where $P^i$ is the $i^{th}$ component of the vector $P^c$.

B. Correctness under DC Load Flow-based CA

In the following we consider DC load flow computation and show that if contingency analysis is performed using DC load flow then the proposed obfuscation algorithm does not affect the result of the contingency analysis, i.e., the error $e_P$ is zero.

The DC load flow model is based on the observation that in a system in normal operation the angular separation along any transmission line is small. This allows one to obtain a linear approximation for (1) of the form

$$P^{DC}_{nm} = V_nV_m(B_{nm}x_{nm}), \quad (9)$$

If one further considers that the per-unit voltages are approximately equal to one, then the power balance equations can be written as

$$\Delta P^{DC} = -P_n + \sum_m B_{nm}x_{nm} = 0. \quad (10)$$

Observe that due to the linearity of the power balance equations in the DC power flow model, the load flow problem for power injection vector $P_l$ can be solved as $x = J^{-1}P_l$.

\textit{Proposition 1:} Under DC load flow based contingency analysis the error introduced through obfuscation $e_P = 0$, where $0$ is the vector of all zeros.

\textit{Proof:} Consider the error $e_P$ introduced by obfuscation in the result of the contingency analysis, as defined in (7),

$$e_P = P^c - \tilde{P}^c$$

$$= P^c - (P^{c,O} - H_cJ_c^{-1}P^{c,O}_i)$$

$$= H_cJ_c^{-1}f^i(Jx + Jx^o) - (H_cJ_c^{-1}f^i(P_l + P^o_i) - H_cJ_c^{-1}f^iP^o_i)$$

$$= H_cJ_c^{-1}f^i(Jx + Jx^o) - H_cJ_c^{-1}f^iP^o_i = 0,$$

where $J_c^{-1}f^i = x^O$ because of (4).

Note that the proof relies on the linearity of the power balance equations in the DC model, which implies that the DC load-flow problem can be solved in one iteration. Thus, the proof does not hold for AC load-flow based contingency analysis.
C. Choosing the Obfuscation Vector

In order to make obfuscation suitable for AC load-flow based contingency analysis, the choice of the obfuscation vector should be such that obfuscation does not introduce a significant error in the result of the contingency analysis, thus obfuscation should not be too big. At the same time, obfuscation should be big enough to hide the actual power flows from an attacker in the following sense. On the one hand, it should be ambiguous for an attacker whether a contingency exists in the actual system in case a critical contingency exists in the obfuscated system. On the other hand, if there is no critical contingency in the obfuscated system, the attacker can be aware of that there is no critical contingency in the actual system either, as this information cannot be used against the system.

The above two requirements imply that the power flow obfuscation $P^\circ$ has to be bounded, and the obfuscation should have maximal entropy. We use the following result from [11] to construct the maximum entropy distribution.

Lemma 1: Fix real numbers $a < b$ and $\mu \in (a, b)$. The continuous probability density function on the interval $[a, b]$ with mean $\mu$ that maximizes entropy among all such densities (on $[a, b]$ with mean $\mu$) is a truncated exponential density

$$q_\alpha(x) = \begin{cases} \frac{C_\alpha e^{\alpha x}}{\int_a^b C_\alpha e^{\alpha x} dx} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

where $C_\alpha$ is chosen so that $\int_a^b C_\alpha e^{\alpha x} dx = 1$, and $\alpha$ is the unique real number such that $\int_a^b x C_\alpha e^{\alpha x} dx = \mu$.

For $\alpha = 0$ the distribution is uniform on $[a, b]$, and its differential entropy is

$$h(X) = \int_a^b q_0(x) \log q_0(x) dx = \log(b - a). \quad (12)$$

Proof: We refer to [11] for the proof. \hfill \blacksquare

As our objective is to obfuscate the power flows, we define the obfuscation vector in terms of the obfuscation power flows $P^\circ$, and use the uniform distribution for obfuscation. We thus define the diagonal matrix $U$ with diagonal elements $U_{i,i} \sim U(0, 0.1)$, and create the vector

$$\hat{P}^\circ = U P. \quad (13)$$

This vector cannot be used directly for the obfuscation because it does not necessarily correspond to any system state. We therefore perform a linearized state estimation on this vector to obtain the state obfuscation vector

$$x^\circ = (H^T H)^{-1} H \hat{P}^\circ, \quad (14)$$

where $H^T$ is the transpose of $H$. Note that the components of $x^\circ$ do not follow a uniform distribution, but the components of the power flow obfuscation vector $P^\circ = H x^\circ$ are likely close to uniform (relative to the actual power flows). Numerical results presented in Section V show that this is indeed the case.

V. PERFORMANCE EVALUATION

In the following we illustrate the efficiency of the proposed algorithm via simulations.

Fig. 2. Histogram of the differential entropy of the relative obfuscation power flows for $u_{max} = 0.1$, computed over 100 runs.

A. Simulation methodology

We used the IEEE 118 bus test system and and used Matpower for the AC load flow based CA. The power flows and injections are represented using p.u, where 1 p.u. equals to 100MW. For the obfuscation, we considered all active power injections and all active power flows, both ‘to’ and ‘from’ buses (hence negative values in the Figures).

B. Obfuscation performance

We first consider the performance of the algorithm in terms of the obfuscation it provides. Note that the level of obfuscation does not depend on the particular contingency considered, it depends on the system topology and the actual system state. These results are thus general for the IEEE 118 bus system.

Figure 2 shows a histogram of the differential entropy of the relative obfuscation power flows, i.e., that of $P^\circ/P$ in a component-wise sense, computed over 100 randomly chosen obfuscation power flows for $u_{max} = 0.1$. We approximated the differential entropy by creating a histogram with 200 bins and using the histogram bins width for numerical integration. The differential entropy of $U(0, 0.1) \approx -3.2$, thus aligning the power flows with the range space of the Jacobian in (14) does alter the distribution of the power flows, but it does not decrease its entropy. In fact, the obfuscation of some power flows by far exceeds $u_{max} = 0.1$, which is the reason for the significantly higher entropy than with the uniform distribution.

Figure 3 shows the QQ plot of the distribution of $P^\circ = H x^\circ$ defined in (14) normalized by $P$, compared to a uniform distribution on $[0, 0.1]$, computed over 100 randomly chosen obfuscation vectors. Recall that the components of $P^\circ$ follow a uniform distribution, but due to (14) the components of $P^\circ$ do not necessarily do so. Figure 3 shows that the distribution of $P^\circ$ indeed differs from uniform, especially at the tails, which is also the reason for the increased differential entropy, as discussed above. At the same time, the body of the distribution is close to uniform. The figure also shows that there is not much difference between the individual obfuscation vectors, as the percentiles are rather close to the average.

These two figures indicate the choice of the obfuscation vector $P^\circ$ provides a good level of randomness thus making it...
compared to the actual power flows. The results shown were obtained for an error in a power flow or a power injection. The results in the corresponding power flows are shown on the horizontal axis, and the errors in the power flows under the considered contingency after deobfuscation: system, thus the scenario corresponds to a severe contingency.

The figure shows that the pre-contingency power flow on transmission line that connects buses 9 and 10. The effect of considering a contingency that affects branch 9 is above 4 p.u., and is the largest power flow in the system, which confirms that obfuscated CA would be viable.

The average RMSE is very small compared to the actual power flows introduced in the result of the CA by obfuscation as a function of the upper bound $u_{\text{max}}$ of the uniform distribution used for obfuscation in (13). The average RMSE is defined as $\|e_p\|_2$, where $|.|$ is the number of components in the vector. For every $u_{\text{max}}$ value the figure shows the mean over 100 simulations together with the 5 and 95 percentiles. The figure shows that the average RMSE increases approximately linearly over a wide range of $u_{\text{max}}$ values, and so do the percentile values. The average RMSE is very small compared to the actual power flows in the system, which confirms that obfuscated CA would be viable.

Figures 7 and 8 show the difference between the obfuscated power flows $P^{c,O}$ and the power flows obtained without obfuscation $P^c$ for two different obfuscation vectors $x^o$, but only for those power flows that increase due to the contingency. The vertical axis is thus effectively the introduced obfuscation. Both figures show that the amount of obfuscation grows with the power flow, but the actual values differ because they depend on the obfuscation vector $x^o$. The two obfuscation vectors used for Figure 7 and for Figure 8 were chosen from the considered 100 obfuscation vectors for $u_{\text{max}} = 0.1$ so as to represent two different scenarios in terms of the signs of the introduced obfuscations per flow. In the first scenario (Figure 7), all power flows that increased due to the contingency without obfuscation

hard for an adversary to guess the real power flows. We now turn to the evaluation of the error introduced by obfuscation for AC load flow-based CA.

1) Obfuscation vs. CA accuracy: In the following we consider a contingency that affects branch 9, which is a transmission line that connects buses 9 and 10. The effect of the contingency on active power flows is shown in Figure 4. The figure shows that the pre-contingency power flow on branch 9 is above 4 p.u., and is the largest power flow in the system, thus the scenario corresponds to a severe contingency.

Figure 5 shows a scatter plot of the error vs. the power flows under contingency obtained with regular CA.

Figure 6 shows a scatter plot of the error vs. the power flows and injections $P^o/P$ for $u_{\text{max}} = 0.1$ vs. a uniform distribution $U(0,0.1)$, computed over 100 runs.
have a positive amount of obfuscation, while in the second scenario (Figure 8), there are a few relatively small power flows for which the obfuscation is negative. Power flows that have a negative obfuscation are determined by the obfuscation vector $x^O$, which is unknown to the attacker. Consequently, by just observing $P^{c,O}$, an attacker cannot be certain how much obfuscation is introduced and for which flows the obfuscation is negative. Thus, the fact that there is a thermal capacity violation in the obfuscated system does not imply that it is also the case after de-obfuscation, and thus an attacker that observes a violating contingency based on $P^{c,O}$ cannot be certain that there is a violating contingency in the actual system, according to $P^c$.

VI. Conclusion

We proposed an approach to obfuscate information regarding power flows to enable CA in the cloud while allowing the operator to obtain accurate post contingency flows. Our approach doesn’t introduce any error for CA using a DC model and our numerical results show that the error introduced when using AC models is tolerable. It is subject of our future work to extend the obfuscation algorithm so that it always introduces positive obfuscation to the power flows that increase due to contingency. Furthermore, our future work will include analytically bounding the error introduced by the proposed obfuscation and an analytical characterization of the randomness of the obfuscation vector.

Acknowledgements

This material is based upon work supported in part by the Department of Energy under Award Number DE-OE0000097.

REFERENCES


