Risk-Aware Vulnerability Analysis of Electric Grids from Attacker's Perspective

Yihai Zhu, Jun Yan, Yan (Lindsay) Sun, and Haibo He
Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881
Email: {yhzhu,jyan,yansun,he}@ele.uri.edu

***Abstract*—Electric grid is one of the largest interconnected networks on the earth, and is vital to the operation of modern society. Within recent decades, the occurrence of several large scale power blackouts raised many concerns from different aspects. For example, the most recent India power blackout in July 2012 affected 620 Million people. Investigating the vulnerability of electric grids becomes increasingly important and urgent. In this paper, we study the vulnerability of electric grids from attacker's point of view. First, the extended model based on DC power flow analysis is adopted to simulate cascading failures in electric grids; then a novel metric, called the *risk graph*, is proposed to reflect the hidden relationship among substations in terms of vulnerability; finally a practical multiple-node attack strategy is developed and proved to be stronger than the traditional load based approach on IEEE 57 and 118 bus systems. This work provided a new point of view toward understanding cascading failures in electric systems.**

***Index Terms*—Cascading failure, Electric grid vulnerability, Security, Attack**

I. Introduction

The security and reliability of electric grids have attracted increasing attentions from different areas after several large-scale blackouts, such as the cases of North American 2003 [1], South American 2009 [2] and India 2012 [3]. Today, many organizations, e.g. IEEE Power and Energy Society (PES), are devoted to developing the methodologies and tools to investigate the vulnerability of electric grids.

Large scale power outages are often referred to as *cascading failures* [4]. In general, the occurrence of them includes four sequential steps. First, one or more components (e.g. substations or transmission lines) partially or completely fail; second, those failed components shift their load to other components nearby; third, the new load of those nearby components are over their capacity, and then those components become overloaded, fail to work and shift their load to other surviving components; finally, the failure propagates from one or a few points to the whole electric grid. Cascading failures usually caused disastrous consequences. To understand cascading failures, an efficient perspective is from the *attack* point of view, where cascading failures could be simulated under different models. In reality, attacks might be triggered by many initial events intentionally, e.g. cyber attacks [5] and terrorist threats [6], or unintentionally, e.g. natural incidents [1].

In the current literature, the vulnerability analysis of electric grids from attack perspective could be conducted under different models, e.g. pure power flow models [4], [7], [8], pure topological models [9]–[12] and hybrid models [13]–[15]. Under *pure power flow models*, the well-known *contingency analysis*, also called $N-x$ *criterion*, is a big family [4], [7]. Those approaches are mainly employed to identify the criticality of nodes/links, the protection of which in advance might inhibit cascading failures. Although single contingency analysis is doable, multiple contingencies are often computationally infeasible, even some approaches are exploited to speed up them [8].

Different from pure power flow models, there is extensive literature on modeling cascading failures by employing abstract network theories [16], called *pure topological models* basically including the *recoverable model* [9] and the *non-recoverable model* [11], [12]. Although pure topological models are hard to correctly represent power distribution in real electric systems, they are still very useful to investigate the vulnerability of electric grids from the attack perspective, especially providing some *metrics* to define stronger attacks. Many metrics have been employed to assist attackers, e.g. *degree* and *load* in [9], *load distribution vector* (LDV) in [10], and *risk if failure* (RIF) in [12]. There is no doubt that the way to obtain stronger attacks is a significant perspective to understand cascading failures in electric grids.

In addition to the previous two kinds of models, *hybrid models* employ both electric features (e.g. impedance and power transmission distribution factors (PTDFs)) and abstract network features (e.g. betweenness). A DC model based hidden failure analysis approach [13] was proposed to investigate the error and attack tolerance of electric grids. Another hybrid model is well extended from the recoverable model [9], and called the *extended model*, which was first proposed in [14] and further developed in [15]. The work in [15] is to rank the criticality of nodes/links of an electric grid following the philosophy of $N-1$ criterion. However, investigating cascading failures under the extended model are promising to understand how cascading failures occur in electric systems. The work presented in this paper is aligned with the direction adopting the extended model to study cascading failures.

In this study, the first goal is to demonstrate how cascading failures occur under the extended model. Although the extended model has been adopted to study the vulnerability of electric grids, it is still incomplete to model cascading failures. Because none of existing literature has demonstrated how cascading failures occur under the extended model.

The second goal is to discuss the *node attack strategy*

(NAS) under the extended model. In this study, the *attack* means the failure of one or more nodes simultaneously; while *node attack strategy* means how to choose *target nodes* (TNs). From the attack performance point of view, the *optimal attack strategy*, enumerating all possible node combinations, could obtain the best attack. However, the optimal approach is often computationally infeasible. Instead, a novel search based approach, called the *sub-optimal attack strategy*, is proposed, which has low computational complexity and good attack performance. However, the sub-optimal approach has its own limitations. In order to develop practical attack strategies, a novel metric, called the *risk graph (RG)*, is proposed to reflect the hidden relationship between nodes. Adopting the risk graph, we propose a novel and fast attack strategy, called the *risk graph based attack strategy*, which is tested and compared with the traditional load based approach on IEEE 57 and 118 bus systems. The simulation results demonstrate the risk graph based approach is an efficient attack strategy under the extended model.

The paper is structured as follows. Cascading failures under the extended model are discussed in Section II. In Section III we describe the risk graph and node attack strategies in detail. Simulations and observations are made in Section IV. Finally, general conclusion is provided in Section V.

II. Cascading Failures under the Extended Model

A. The Extended Model

In this paper, we consider an electric grid as a directed graph $\mathbf{G} = \{B, L\}$, where B and L are nodes (i.e. substations) and links (i.e. transmission lines) sets, respectively. All generators and all load substations (the substations that delivery power to customers) are denoted as sets G and D, respectively, where $G \subseteq B$ and $D \subseteq B$. In addition, N_B, N_L, N_G and N_D are adopted to represent the number of nodes, links, generators and load substations, respectively.

We adopt the extended model to simulate the power distribution in electric grids. The extended model was extended from the recoverable model by employing electric features. Generally speaking, it has four major differences from the recoverable model: (1) the linear model analysis, (2) the extended betweenness, (3) the electric distance and (4) the net-ability. We will briefly introduce the extended model by representing the first two different concepts, which are related to our work, as follows. The details of the extended model were discussed in [14], [15].

1) *Linear Model Analysis*: The power distribution in real electric systems follows electric theories, e.g. Kirchhoff's law. *Power Transfer Distribution Factors* (PTDFs) are computed based on DC power flow [17]. PTDFs reflect the sensitivities of power flow changes in transmission lines due to the nodal real power injections. The basis of the extended model is to employ PTDFs to represent the power distribution in electric systems, which is absolutely different from adopting the shortest paths under the recoverable model [9]. In this paper, all PTDFs are calculated by using MATPOWER [17], a Matlab-based tool for solving power flow analysis problems.
2) *Extended Betweenness*: In electric grids, power is transmitted from generators to load substations. The power flow change in each transmission line caused by each generator-load substation pair could be calculated from PTDFs matrix [15]. In other words, the accumulation of all power on a specific transmission line caused by all generator-load substation pairs would determine the total power on this link. The extended betweenness of a node, e.g. node i, is defined as half of the total summation of power flow in the links connecting to node i. The extended betweenness of a node is adopted as its *load*, similar to the functionality of betweenness in [9].

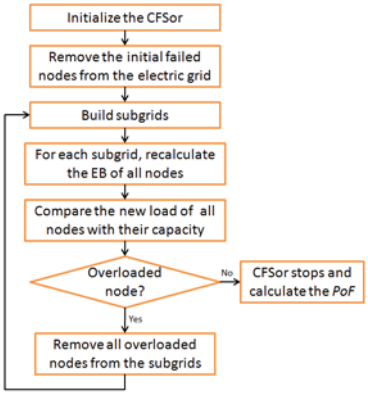

Fig. 1. Algorithm diagram of the cascading failure simulator (CFSor) under the extended model used in this paper.

B. Cascading Failure Simulator (CFSor)

Adopting the extended model to study cascading failures, we redefine the concepts of the *load*, *system tolerance* and *capacity*, which are the basic concepts under pure topological models [9], [11]. First, the load of a node is defined as its extended betweenness; Second, the system tolerance is denoted as α, and has a range from 1.05 to 2 with an interval 0.05; finally, the capacity of a node is defined as the multiplication of its initial load with the system tolerance. In addition, we adopt the *percent of failure* (PoF) [12] as the measurement metric to evaluate the damage,

$$PoF = 1 - \frac{N_B{}'}{N_B} \tag{1}$$

where N_B and $N_B^{'}$ are the number of surviving nodes in an electric grid before and after an attack.

Cascading failures under the extended model are conducted by using the cascading failure simulator (CFSor) illustrated in Fig. 1. The CFSor has three major parts. First, initialize the CFSor and trigger the initial failures. Second, conduct the failure propagation until no overloading node in the electric grid. Finally, measure the damage after the failure stops.

III. Risk-Graph and Node Attack Strategy

Electric substations are the major components in electric grids. Their security and safety are strongly related to the reli-

TABLE I
THE TOP TEN STRONGEST NODE COMBINATIONS FOR NSA^M_{opt}, WHERE M=1,2,3 AND α = 1.2, ON IEEE 118 BUS SYSTEM.

Index	NSA^1_{opt}		NSA^2_{opt}		NSA^3_{opt}	
	TNs	PoF (%)	*TNs*	PoF (%)	*TNs*	PoF (%)
1	70	40.7	**17**,38	49.2	**17,38**,94	58.5
2	23	38.1	**70**,98	49.2	**17,38**,96	57.6
3	38	36.4	**38**,69	47.5	**38,69**,94	57.6
4	65	34,8	**38**,94	47.5	**17,38**,69	56.8
5	24	33.9	**23**,98	46.6	**30,38**,94	56.8
6	19	30.5	**38**,96	46.6	**17,38**,82	55.9
7	34	27.1	**30**,38	45.8	**30,38**,69	55.9
8	68	27.1	**30**,65	45.8	**30,38**,96	55.9
9	30	25.4	**70**,86	45.8	**38,69**,96	55.9
10	17	16.1	**70**,89	45.8	70,88,98	55.1

ability of electric grids. In reality, substations are confronting various risks, e.g. natural disasters (the fall of trees [1]), cyber attack [5], and so forth. Hence, investigating the cascading failures from substations perspective is an urgent task.

Under the extended model, substations are considered as nodes. We will investigate the vulnerability of nodes by discussing *Node Attack Strategy (NAS)*. In the context of studying cascading failures, an attacker's goal is to identify a set of *target nodes* (TNs), whose simultaneous failure causes large PoF to an electric grid.

The *traditional load based NAS* is well studied under pure topological models and shown to be a strong approach [9], [11]. We will adopt this approach as a reference approach. Under the extended model, if an attacker wants to choose the M target nodes, the load-based NAS, denoted as NAS^M_{load}, works as follows,

* NAS^M_{load}: The nodes with the top M largest load will be chosen as target nodes.

A. Sub-optimal Node Attack Strategy

For an attacker, the strongest NAS is no doubt the exhaustive search, also called as the *optimal NAS* and denoted as NAS^M_{opt},

* NAS^M_{opt}: Choose the M nodes, whose simultaneous failure yields the largest PoF under a given system tolerance (α), as the target nodes.

Although the NAS^M_{opt} yields the best attack, it is very time-consuming and mostly computationally feasible. Take IEEE 118 bus system as example. If we launch five-node attack ($M = 5$), there are $\frac{118!}{5!(118-5)!} = 174,963,438$ candidate combinations for NAS^5_{opt}. In addition, the time of calculating PTDFs of IEEE 118 bus system once needs an average 0.01 second by using Matlab under Window 7 OS with 4 GB memory and dual-core i5 CPU (2.4GHz each), which means only computing the PTDFs of all combinations approximately needs 20 days.

Although NAS^M_{opt} is mostly computationally infeasible, it is still doable when M and N_B are proper. For example, if the size of a power grid network, N_B, is more than moderate, such as $N_B \geq 300$, the number of target node, M, should be small, such as $M \leq 4$. We conducted the experiments by using NAS^M_{opt} on IEEE 118 bus system, where M is set to be $1, 2, 3$ and α is set to be 1.2. In Table I, the top ten strongest node combinations are shown. An interesting observation is made.

Procedure 1 Obtain the *round recommended combination set* (RRCS) and choose target nodes (TNs) for NAS^M_{subopt}.

1: Give a system tolerance (e.g. $\alpha = 1.2$) and the number of TNs (i.e. M).
2: //Choose the TN for NAS^1_{subopt} or do the initialization of the iteration.
3: **for** i = 1 : N_B **do**
4: Using the CFSor, launch one-node attack by removing node i under given α, then record the PoF caused by this attack.
5: **end for**
6: **if** M == 1 **then**
7: Choose the node with the largest PoF as the TN for NAS^1_{subopt}.
8: **Goto the last step of the procedure**.
9: **else**
10: Choose the nodes with the top P largest PoF as the candidate nodes and put them into S_C. Meanwhile, choose the nodes with the top R largest PoF as the 1^{th} RRCS and put them into S^1_{RRC}
11: **end if**
12: //Start the iteration, and obtain S^m_{RRC} under given S^{m-1}_{RRC}.
13: **for** m = 2 : M **do**
14: **for** i = 1 : R **do**
15: Get the i^{th} combination in S^{m-1}_{RRC}, denoted as $\mathbf{C}_i$.
16: **for** j = 1 : P **do**
17: Get the j^{th} candidate node in S_C, denoted as n_j.
18: Combine $\mathbf{C}_i$ and n_j to get a new candidate combination.
19: **end for**
20: **end for**
21: Using the CFSor, conduct multi-node attack for each new candidate combination ($R \times P$ in total), and record all PoF.
22: Choose the new combinations with the top R largest PoF as the m^{th} RRCS, and put them into S^m_{RRC}.
23: **end for**
24: The node combination in S^M_{RRC}, which can cause the largest PoF, are the TNs for NAS^M_{subopt}.
25: The procedure ends

In Table I, the highlighted node or node combinations illustrate that the top strongest combinations for NAS^m_{opt} ($m \geq 2$) are mainly obtained by joining some of the top strongest combinations for NAS^{m-1}_{opt} with another critical node. For instance, the first node in the node combinations for NAS^2_{opt} (i.e. node 17, 70, 38, 23 and 30) are all from the top ten strongest nodes for NAS^1_{opt}; the first two nodes of the node combinations for NAS^3_{opt} (i.e. node combinations $\{17, 38\}$, $\{38, 69\}$, $\{30, 38\}$ and $\{70, 88\}$) are almost from the top ten node combinations for NAS^2_{opt}, except $\{70, 88\}$, which occurs just once. This observation is reasonable. If a node combination for NAS^{M-1}_{opt} could cause large PoF, the new node combination by adding another critical candidate node to this node combination probably yields larger PoF for NAS^M_{opt}, though the new node combination might not be the strongest one for NAS^M_{opt}.

Inspired by the above discussions, a novel sub-optimal search based NAS, denoted as NAS^M_{subopt}, is proposed. Before discussing the algorithm of NAS^M_{subopt}, we have three statements. First, the algorithm procedure mainly consists of M rounds, including one initial round and $M-1$ iterative rounds. Second, in each round (e.g. $m^{th}round$), the top strongest node combinations are chosen as the *round recommended combination set* (RRCS), denoted as S^m_{RRC}. Finally, there are two important parameters, P and R, in this algorithm, which are adopted to control the initial size of candidate node set, denoted as as S_C, and the size of RRCS, respectively.

Suppose an attacker wants to launch attacks by adopting NAS^M_{subopt}, Procedure 1 shows the steps to obtain target nodes for NAS^M_{subopt}. From Procedure 1, we know that NAS^M_{subopt}

TABLE II
AN REALIZATION OF THE RECOMMENDED COMBINATION SET ON IEEE 118 BUS SYSTEM.

Index	S^1_{RRC}	S^2_{RRC}	S^3_{RRC}	S^4_{RRC}	S^5_{RRC}	S^6_{RRC}	S^7_{RRC}	S^8_{RRC}
1	70	17,38	38,17,94	38,17,94,69	38,17,94,69,103	38,69,94,30,103,7	38,69,94,30,103,7,98	38,69,94,30,103,7,98,99
2	23	70,98	38,17,96	38,17,96,69	38,17,96,69,103	38,17,94,69,103,98	38,69,94,30,103,7,33	38,69,94,30,103,11,98,99
3	38	38,69	38,69,94	38,69,94,30	38,69,94,30,103	38,17,94,69,103,99	38,17,94,69,103,98,99	38,69,94,30,103,11,98,33
4	65	38,94	38,17,69	38,69,30,96	38,17,94,69,98	38,17,94,69,103,33	38,17,94,69,103,98,33	38,69,94,30,103,11,99,33
5	24	23,98	38,94,30	38,17,69,82	38,17,94,30,7	38,69,94,30,103,11	38,17,94,69,103,99,33	38,69,94,30,103,7,98,50
6	19	38,96	38,17,82	38,17,69,103	38,69,30,96,103	38,69,30,96,103,7	38,69,94,30,103,11,98	38,69,94,30,103,7,98,47
7	34	70,89	38,69,30	38,17,94,103	38,17,94,69,106	38,17,94,69,103,96	38,69,94,30,103,11,99	38,69,94,30,103,7,98,99
8	68	70,86	38,69,96	38,17,94,66	38,17,94,69,33	38,17,94,69,103,29	38,69,94,30,103,11,33	38,69,94,30,103,7,98,87
9	30	70,112	38,96,30	38,17,96,103	38,17,94,69,117	38,17,94,69,103,31	38,69,94,30,103,7,96	38,69,94,30,103,7,98,93
10	17	70,116	70,98,88	38,69,94,26	38,17,94,69,98	38,17,94,69,103,50	38,69,94,30,103,7,50	38,69,94,30,103,7,98,95
11	31	30,38	38,17,83	38,69,94,5	38,69,94,30,11	38,17,94,69,103,16	38,69,94,30,103,7,63	38,69,94,30,103,7,98,97
12	80	30,65	38,69,82	38,17,69,83	38,69,30,96,7	38,17,94,69,103,47	38,69,94,30,103,7,47	38,69,94,30,103,7,33,96
13	64	70,16	65,30,94	38,17,69,92	38,17,94,69,29	38,17,94,69,103,113	38,69,94,30,103,7,99	38,69,94,30,103,7,33,50
14	61	70,74	65,30,96	38,69,30,82	38,17,94,69,16	38,17,94,69,103,9	38,69,94,30,103,7,13	38,69,94,30,103,7,33,99
15	37	70,91	70,98,93	38,17,96,66	38,17,94,69,47	38,17,94,69,103,10	38,69,94,30,103,7,35	38,69,94,30,103,7,33,86
16	69	38,82	70,98,95	38,69,94,25	38,17,94,69,9	38,17,94,69,103,18	38,69,94,30,103,7,86	38,69,94,30,103,7,33,87

needs to do $P\times R\times(M-1)+N_B$ cascading failure simulations to obtain its best attack. Theoretically, the computational complexity of NAS^M_{subopt} is $P \times R \times (M-1) + N_B$. When investigating multi-node attack, the computational complexity of NAS^M_{subopt}, at the worst case ($P = R = N_B$), is appropriate to $O(M(N_B)^2)$, which is much lower than $O((N_B)^M)$ of NAS^M_{opt}.

B. The Introduction of Risk Graph

Although the sub-optimal NAS could sharply reduce the computational complexity of the search based approaches, it still has two limitations. First, the sub-optimal NAS is not suitable for real-time attack, because it still needs lots of time to do search, especially when N_B is large. Second, NAS^M_{subopt} needs to estimate the system tolerances before attacks, which is nearly impossible in reality due to many reasons. In this subsection, we propose a novel metric, called the *risk graph* (RG), which can yield better attack strategy.

The risk graph is inspired by the *round recommended combination set* (RRCS). One realization of the RRCS is shown in Table II, from which we make an insightful observation that there are some fixed patterns of the occurrence of certain nodes or node combinations, e.g. node 38 and node combination $\{38, 17\}$. Studying this hidden relationship between candidate nodes may help people to understand the vulnerability of electric grid networks.

An risk graph is constructed as follows,

1) First, given α, M, R and P, conduct the sub-optimal approach in Procedure 1, and obtain the intermediate results, the round recommend combination set (RRCS) represented as $S^1_{RRC}, ..., S^M_{RRC}$.
2) Check the RRCS (An example is shown in Table II). If a node occurs in the RRCS, this node becomes a vertex of the risk graph. In addition, each vertex has a *vertex occurrence frequency (VOF)*, defined as the number of the corresponding node appears in the RRCS.
3) Add an edge between each pair of vertexes and set its initial weight to be zero. The weight of an edge is referred to as the *edge occurrence frequency* (EOF).
4) Examine the RRCS and update the EOF of all edges. Suppose a pair of nodes, say node i and node j, appears in a combination that has m nodes, increase the EOF of the $edge_{i-j}$ by adding $\frac{2}{m(m-1)}$. For example, for the combination $\{38, 17, 94\}$, the EOF of three edges, $edge_{38-17}$, $edge_{38-94}$ and $edge_{38-94}$, are increased by 1/3. As another example, assume node 38 and node 17 appear simultaneously in three node combinations: $\{38, 17\}$, $\{38, 17, 94\}$ and $\{38, 17, 94, 69\}$, the EOF of $edge_{38-17}$ is $1 + 1/3 + 1/6 = 3/2$.
5) Finally, delete the edges with EOF as zero and the vertexes that are sole.

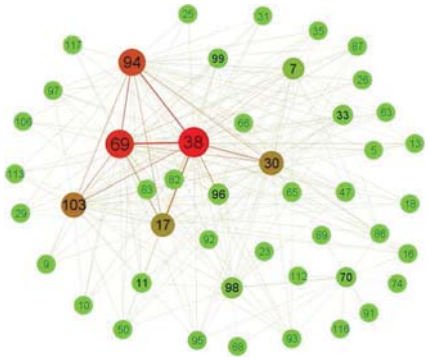

Fig. 2. The risk graph on IEEE 118 bus system, where the system tolerance, α, is set to be 1.2. This figure is visualized by Gephi [18].

The risk graph based on Table II is shown in Fig. 2, generated by *Gephi* [18]. In the risk graph, the size and color of a vertex is decided by its VOF. And the width and color of a edge is determined by its EOF. The bigger (wider) and redder a vertex (or edge) is, the larger its VOF (or EOF) is.

C. Risk-graph Based Node Attack Strategy

The construction of a risk graph is mainly affected by two factors, the system tolerance (α) and the parameters (P and R) of NAS^M_{subopt}. The former is major factor, and the later is minor factor. That is, the risk graph under a certain α is probably different from under another α. In order to obtain a robust risk graph, we first generate the single risk graphes under different system tolerances, and then average those single risk graphs to obtain the *average risk graph* (ARG). In this study, an ARG is obtained by two steps. First, generate single risk graphs under α from 1.05 to 2 with an interval 0.05 (There are twenty α in total.). Then, average those twenty single risk graphs, where the "averaging" means (1) the VOFs

(EOFs) of a vertex (an edge) appearing in more than one single risk graphs are added together, (2) divide the new VOFs (EOFs) by the number of single risk graphs (20 in this study). The figure of an ARG is similar to that in Fig. 2.

The ARG of an electric grid network has two advantages, robust to the system tolerance and reflecting the hidden combination relationship among candidate nodes. Those advantages are helpful to find practical attack strategies. In this study, we propose a novel attack strategy, called the *risk graph based attack strategy*, based on the ARG of an electric grid. Suppose an attacker has already obtained the ARG of an electric grid, the risk graph based attack strategy, represented as $NAS^M_{riskgraph}$, is conducted as follows,

* $NAS^M_{riskgraph}$: When $M = 1$, choose the node with the largest VOF as the target node for $NAS^1_{riskgraph}$. Otherwise, the M target nodes for $NAS^M_{riskgraph}$ are chosen from the ARG by meeting with two restrictions. First, each pair of nodes has a direct edge between them, which means there are $\frac{M(M-1)}{2}$ edges between those M nodes. Second, the summation of all EOFs of those $\frac{M(M-1)}{2}$ edges is maximum.

In this section, we mainly discuss the motivation and construction of the risk graph. The comparisons among the attack strategies mentioned above are made in Section IV.

IV. Simulation Results

We used Matlab to implement the simulations, including modeling the extended model, modeling cascading failures and attacks. The proposed attack strategies are tested and compared with other approaches on IEEE 57 and 118 bus systems [19]. The observations and discussions are made in detail in the following two subsections.

A. Simulation Results for the Optimal and Sub-optimal Attack Strategies

In this subsection, the comparison between the optimal and the proposed sub-optimal approaches are made. The two approaches are tested on IEEE 57 bus systems. Fig. 3 shows the comparison, in which the horizontal axis and the vertical axis represent the the number of target node and the percent-of-failure, respectively. In addition, the solid blue-pentagram curve represents NAS^M_{opt} and the dashdot red-plus curve represents NAS^M_{subopt}. Due to computational complexity of NAS^M_{opt}, M is set to be less than 5 for NAS^M_{opt}. The M for NAS^M_{subopt} is set to be less than 8. Meanwhile, α is set to be 1.2 for both approaches.

Two observations are made from Fig. 3. First, the attack performance of NAS^M_{subopt} can compete with that of NAS^M_{opt}, especially when M is small. In Fig. 3, the attack performance of NAS^M_{subopt} is exactly equal to that of NAS^M_{opt}, when $M = 1, 2$, and just a little weaker at $M = 3, 4, 5$. When $M > 5$, form the computational complexity comparison made in subsection III-A, we know NAS^M_{opt} is computationally infeasible, but the proposed approach, NAS^M_{subopt}, is doable and could obtain good attack performance. Second, as M increases, the curves in Fig. 3 go flat and probably arrive at

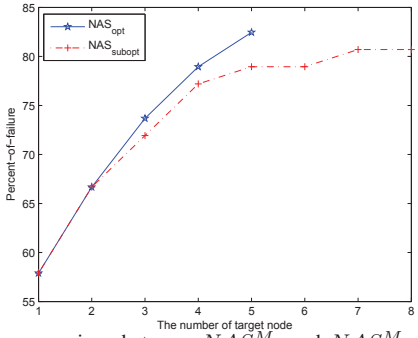

Fig. 3. The comparison between NAS^M_{opt} and NAS^M_{subopt} on IEEE 57 bus system, where the system tolerance (α) is set to be 1.2.

their upper bounds. Thus, studying cascading failures at small M values will be more meaningful than at large M values.

The comparison is also conducted on IEEE 118 bus system, and the observation is similar to that in Fig. 3. Due to space limitation, we do not show that figure in this study.

In summary, the attack performance of the sub-optimal approach can compete with that of the optimal approach, and the computational complexity of NAS^M_{subopt} is much lower than that of NAS^M_{opt}. Thus, NAS^M_{subopt} can substitute NAS^M_{opt}, when the optimal approach is unreachable.

B. Comparison among Different Attack Strategies

In this subsection, the proposed sub-optimal and the risk graph based approaches are compared with the traditional load based approach. IEEE 57 and 118 bus systems are adopted as the testing data, and the results are shown in Fig. 4 and Fig. 5, respectively. Within those two figures, the horizontal axis represents the number of target node, and the vertical axis represents the percent-of-failure. In addition, the dashdot red-plus curve, solid green-star curve and solid magenta-square curve represent NAS^M_{subopt}, $NAS^M_{riskgraph}$ and NAS^M_{load}, respectively. From Figs. 4 and 5, we make the following observations and discussions.

First, from the attack performance point of view, $NAS^M_{riskgraph}$ is very close to NAS^M_{subopt}, but much stronger than NAS^M_{load}. In Figs. 4 and 5, the curves representing the attack performance of $NAS^M_{riskgraph}$ is pretty close to the curves representing NAS^M_{subopt}, especially when $M \leq 5$. Meanwhile, the curves representing NAS^M_{load} are much lower than previous two types of curves.

Second, as M increases, a sharp drop in the curves representing the performance of both $NAS^M_{riskgraph}$ and NAS^M_{load} occurs. The reasons why the drop happens are different. The load based approach dose not concern the speciality of the cascading failures under the extended model. The specialty is that cascading failures will quickly stop when the whole electric grid is broken into more than one balanced subgrids. The target nodes chosen by NAS^M_{load} usually have higher load, knocking down which might cause the cascading failure procedure to stop just after the a few rounds, which can not cause serious damage to the power grid referring to the PoF. Different from the load based approach, the risk graph based approach choose its target nodes from the ARG. As we know, the ARG can perfectly reflect the combination between a pair

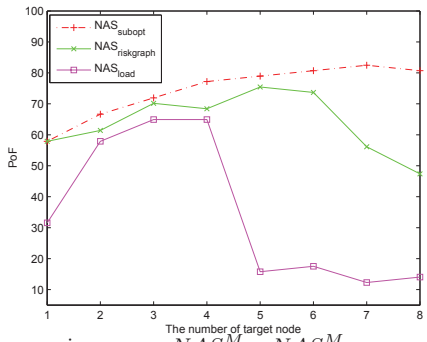

Fig. 4. The comparison among NAS^M_{load}, $NAS^M_{riskgraph}$ and NAS^M_{subopt} on IEEE 57 bus system, where the system tolerance α is set to be 1.2.

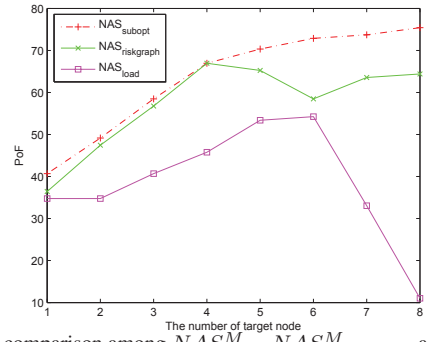

Fig. 5. The comparison among NAS^M_{load}, $NAS^M_{riskgraph}$ and NAS^M_{subopt} on IEEE 118 bus system, where the system tolerance α is set to be 1.2.

of nodes, but can not show the strongest combinations when M is large (e.g. $M > 3$). Thus, the sudden drop in the curves representing $NAS^M_{riskgraph}$ may occur at large M.

Finally, the risk graph based attack strategy is a good choice to launch real-time attacks to electric grids. From the computational complexity perspective, when a ARG are done, which can be constructed off-line, the computational complexity of $NAS^M_{riskgraph}$ is $O(1)$, which is much lower than $O(M(N_B)^2)$ of NAS^M_{subopt} and close to $O(1)$ of NAS^M_{load}. From the attack performance perspective, $NAS^M_{riskgraph}$ is very close to NAS^M_{opt} at small M values and a little weaker at larger M values. However, $NAS^M_{riskgraph}$ is much stronger than NAS^M_{load} referring to POF.

V. Conclusion

In this paper, we adopt the extended model to investigate the cascading failures in electric grids. We first propose a search based sub-optimal attack strategy and use it as the substitution of the optimal attack strategy. Then, a novel metric, called the risk graph, is proposed to show the hidden relationship of nodes. Finally, a novel risk graph based attack strategy is proposed. We compare the proposed approaches with other ones on IEEE 57 and 118 bus systems, and conclude that the risk graph based approach is a good choice to launch real-time attacks under the extended model.

There are three important future directions along this topic. First, link failures are more frequent than node failures in practice. Adopting the extended model to study link failures or joint link/node failures will be desirable to study the vulnerability of electric grids. Second, studying the vulnerability of the large-scale power grids, e.g. the entire North America Electrical Infrastructure data, will be more meaningful. Finally, visualizing the cascading procedures will help people to understand how cascading failures propagate in electric grids.

Acknowledgment

This work is partially supported by NSF award #1117314 and #0643532.

References

[1] U.S.-Canada Power System Outage Task Force, "Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations," April 2004.

[2] Brazil blackout raises more questions for the olympics. TIME. [Online]. Available: http://www.time.com/time/world/article/0,8599,1938011,00.html?xid=rss-topstories

[3] India blackouts leave 700 million without power. The Guardian. [Online]. Available: http://www.guardian.co.uk/world/2012/jul/31/india-blackout-electricity-power-cuts

[4] M. V. (Lead), K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Transactions on Power Systems*, vol. 27, no. 2, 2012.

[5] G. Hug and J. A. Giampapa, "Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.

[6] A. Delgadillo, J. Arroyo, and N. Alguacil, "Analysis of electric grid interdiction with line switching," *IEEE Transactions on Power Systems*, vol. 25, pp. 633–641, 2010.

[7] K. Lo, L. Peng, J. Macqueen, A. Ekwue, and D. Cheng, "Fast real power contingency ranking using a counterpropagation network," *IEEE Transactions on Power Systems*, vol. 13, pp. 1259–1264, 1998.

[8] M. J. Eppstein and P. D. H. Hines, "A "random chemistry" algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Transactions on Power Systems*, vol. 27, no. 3, 2012.

[9] R. Kinney, P. Crucitti, R. Albert, and V. Latora, "Modeling cascading failures in the north american power grid," *Eur. Phys. J. B*, vol. 46, pp. 101–107, 2005.

[10] Y. Zhu, Y. Sun, and H. He, "Load distribution vector based attack strategies against power grid systems," in *accepted by IEEE Global Telecommunications Conference*, Anaheim, CA, USA, Dec.3-7 2012.

[11] J. Wang, L. Rong, L. Zhang, and Z. Zhang, "Attack vulnerability of scale-free networks due to cascading failures," *Physica A*, vol. 387, p. 6671C6678, 2008.

[12] W. Wang, Q. Cai, Y. Sun, and H. He, "Risk-aware attacks and catastrophic cascading failures in u.s. power grid," in *IEEE Global Telecommunications Conference*, 2011, pp. 1–6.

[13] G. Chen, Z. Dong, D. J. Hill, G. H. Zhang, and K. Hua, "Attack structural vulnerability of power grids: A hybrid approach based on complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 389, pp. 595–603, 2010.

[14] S. Arianos, E. Bompard, A. Carbone, and F. Xue, "Powergrid vulnerability: a complex network approach," *EChaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, 2009.

[15] E. Bompard, D. Wu, and F. Xue, "Structural vulnerability of power systems: A topological approach," *Electrical Power Systems Research*, vol. 81, pp. 1334–1340, 2011.

[16] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[17] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower: Steady-state operations,planning, and analysis tools for power systems research and education," *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 12–19, 2011.

[18] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proceedings of the Third International ICWSM Conference*, 2009.

[19] "Power systems test case archive," University of Washington. [Online]. Available: http://www.ee.washington.edu/research/pstca/