Finding critical regions in a network

Stojan Trajanovski, Fernando A. Kuipers, Piet Van Mieghem
Delft University of Technology, The Netherlands
{S.Trajanovski, F.A.Kuipers, P.F.A.VanMieghem}@tudelft.nl

Abstract—It is important that our vital networks (e.g., infrastructures) are robust to more than single-link failures. Failures might for instance affect a part of the network that resides in a certain geographical region. In this paper, considering networks embedded in a two-dimensional plane, we study the problem of finding a critical region - that is, a part of the network that can be enclosed by a given elementary figure (a circle, ellipse, rectangle, square, or equilateral triangle) with a predetermined size - whose removal would lead to the highest network disruption. We determine that there is a polynomial number of non-trivial positions for such a figure that need to be considered and, subsequently, we propose a polynomial-time algorithm for the problem. Simulations on realistic networks illustrate that different figures with equal area result in different critical regions in a network.

Index Terms—geographical failures, critical regions, network robustness, computational geometry

I. INTRODUCTION

Link and node failures in vital infrastructures, such as communication, power grid [1], transportation or mobile [2] networks may be caused unintentionally, for instance due to (aged) equipment failure, power failure [1], or natural disasters [3], or intentionally [4], [5], for example by terrorist attacks or cyber criminals. Attacks or natural disasters often affect a certain geographical area. For instance, devastations from the 2012 catastrophic hurricane “Sandy” in the US stretched from the East Coast to the Lake Area [3]. Similarly, electromagnetic pulses (EMP) might be more stretched than circular in nature [6]. Consequently, failure areas cannot always be sufficiently accurately approximated by circular shapes, as often considered in papers on regional failures (e.g., [7]) and require taking into consideration more two-dimensional figures, such as ellipses, rectangles, and triangles.

We use the term critical region to denote the position of a predetermined two-dimensional figure, for which the failure of all nodes and their incident links covered by that figure will affect the performance of the network most. In this paper, we study the problem of finding critical regions of various shapes in a network.

Our main contributions can be summarized as follows:

* We prove that only a polynomially bounded number of figure positions (for circles, ellipses, squares, rectangles, and equilateral triangles) need to be considered for finding a critical region;
* We propose a polynomial-time algorithm for detecting critical regions for the aforementioned figures;
* We determine the impact of a critical region failure in real-world networks for different figures of equal area.

In general, the level to which connectivity can be maintained under failures has typically been used as the main metric to characterize network robustness (see [8] for an overview). An other way of robustness characterization is by employing probabilistic graph percolation theory [9]. Robustness of networks against geographical circular failures has been studied by Neumayer et al. [7], [10]. Regarding network flow problems, Sen et al. [11] have studied region-disjoint paths constrained by fixed and predetermined critical regions. While most papers confine to the circular failure model, in this paper we consider several types of two-dimensional figures to represent a regional failure.

The remainder of this paper is organized as follows. Our formal model and problem statement are defined in Section II. In Section III we reduce the, in principle, infinite size of the search space of possible locations for the figures to a search space of polynomial size and provide an accompanying algorithm for detecting critical regions. Section IV identifies and studies critical regions in real-world networks. Concluding remarks are given in Section V.

II. MODEL AND PROBLEM STATEMENT

We start with a presentation of our network model and the problem considered.

Model: We represent a network as a weighted (directed or undirected) graph $G(N, \mathcal{L})$ in a plane consisting of a set $N$ of $N$ nodes and a set $\mathcal{L}$ of $L$ links. Each node $i \in N$ has two-dimensional coordinates $(x_i, y_i)$. The Euclidean distance between two nodes $u$ and $v$ is denoted by $d(u,v)$. The weight of a link $(i,j) \in \mathcal{L}$ is denoted by $w(i,j)$. The weight may reflect the distance, but it could also reflect another metric such as link capacity.

We define the critical region $C(F, X)$ to be a region covered by the position of the figure $F$ in the two-dimensional plane for which the removal of all nodes in that area, and the links incident to them, leads to a maximum deterioration in a certain network metric $X$. The network metric $X$ could for instance represent the number of affected nodes, the average shortest path length, the number of connected node pairs, the size of the giant component, or some service function like packet loss or average delay. There might be multiple critical regions that affect the metric $X$ to the same degree.

We will consider several figures as shown in Fig. 1, including the circle $F_C(O(x_j, y_j), \varphi, r)$ with radius $r$, ellipse $F_E(O(x_j, y_j), \varphi, a, b)$ with semi-axes lengths $a$ and $b$, square $F_S(O(x_j, y_j), \varphi, a, b)$ with side length $a$, rectangle $F_R(O(x_j, y_j), \varphi, a, b)$ with side lengths $a$ and $b$, and the
equilateral triangle \( \mathcal{F}_T(O(x_j, y_j), \varphi, a) \) with side length \( a \). Indeed, the ellipse with equal axes is a circle and a rectangle with equal sides is a square. In the remainder, we use the term dimensions to refer to radii or sides in general.

To characterize the position of a figure, we should determine its center \( O(x_j, y_j) \) and orientation \( \varphi \), which is the angle \((0 \leq \varphi \leq 2\pi)\) between the \( x \)-axis and an axis of symmetry in the figure as shown in Fig. 1. The orientation of a circle \( \mathcal{F}_C(O(x_j, y_j), \varphi, r) \) is irrelevant as any rotation applies.

**Critical region problem:** For a given network \( G(N, \mathcal{L}) \) embedded in a plane, find a critical region \( C(\mathcal{F}, X) \) with respect to network metric \( X \) and two-dimensional figure \( \mathcal{F}_{fig}(O(x_j, y_j), \varphi, \text{dim}) \), where \( \text{dim} \) is the vector of dimensions that defines \( \mathcal{F}_{fig} \), \( fig \in \{ \mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{R}, \mathcal{T} \} \).

**III. CRITICAL REGION DETECTION**

We first demonstrate that finding critical regions of a given two-dimensional figure is polynomially solvable for the figures in Fig. 1.

**A. Theoretical basis**

In this section, we use three kinds of geometric transformations defined in Definition 1.

**Definition 1:** Translation of a figure is moving it in parallel to a given line (e.g., the \( x \)-axis in Fig. 2a). Rotation of a figure \((12)\) along a given node assures that the distance between a point on the perimeter of the figure and that node remains the same (Fig. 2a). Finally, we define sliding along two nodes as moving the position of the figure such that these two nodes still lie on the perimeter of the figure. The sliding differs for different figures as visualized in Fig. 2.

**Theorem 1:** For the five figures given in Fig. 1, if there exists a two-dimensional figure that covers a set of nodes \( S \), then that same set \( S \) could also be covered by the same type of figure that passes through at most 3 nodes.

**Proof:** We start from an arbitrary two-dimensional figure \( \mathcal{F} \). We will consider three positions for this figure, denoted - for ease of notation - by \( \mathcal{F}^{(i)} \), for \( i = 0, 1, 2, 3 \), with \( \mathcal{F}^{(0)} \) the initial position of the figure in which it covers all the points in \( S \). If there is no node that lies on the perimeter of \( \mathcal{F}^{(0)} \), then one can translate \( \mathcal{F}^{(0)} \) parallel to the \( x \)-axis until (at least) one node \( A \in S \) hits the perimeter as exemplified in Fig. 2a. We denote the position of this figure by \( \mathcal{F}^{(1)} \) and it still contains all the nodes in \( S \). If \( A \) is the only node in \( S \) on the perimeter of \( \mathcal{F}^{(1)} \), we rotate, either clockwise-wise or counter-clock-wise until (at least) one node \( B \in S \), different from \( A \), hits the perimeter as shown in Fig. 2a. Denoting the position of this figure by \( \mathcal{F}^{(2)} \), one can slide \( \mathcal{F}^{(2)} \) until at least one more node \( C \in S \) lies on the perimeter of \( \mathcal{F}^{(2)} \). Sliding is always possible for the considered figures, as illustrated in Fig. 2. The resulting position \( \mathcal{F}^{(3)} \) still contains all nodes in \( S \) and is characterized by (at most) three nodes. Sliding of
If there is a set of collinear nodes (even be examined by a figure through the non-collinear node and included. If a figure covers some set of collinear nodes and at
In Theorem 2 collinear or “quasi collinear” nodes were excluded. Three nodes are “quasi collinear” if and only if one of them is on a distance $a$ or $b$ reflecting the dimensions of the figure) from the line through the two other nodes.

**Theorem 2:** The positions of critical ellipses, rectangles and equilateral triangles can be uniquely characterized by three nodes, unless the nodes are collinear for the triangle and the rectangle or “quasi collinear” for the rectangle.

**Proof:** The proof relies on the figure equations from analytic geometry [12], [13] and appropriate case analysis and can be found in Appendix A. Crucial possibilities are visualized in Fig. 3.

In Theorem 2 collinear or “quasi collinear” nodes were excluded. If a figure covers some set of collinear nodes and at least one node that is not collinear to them, this case will be examined by a figure through the non-collinear node and two other nodes. If there is a set of collinear nodes (even with cardinality greater than 3) that could be covered by some position, but no other non-collinear node can form a figure with them, then it is enough to consider two crucial positions: where one of the two “end nodes” lies in a corner of the figure and all the other nodes are positioned on a single side (Fig. 3c). Similarly, for “quasi collinear” nodes, two crucial positions have to be examined (Fig. 3d). Fortunately, these cases are special instances for when two nodes lie: (i) on a same side for collinear (Fig. 3a) or (ii) on parallel sides for “quasi collinear” (Fig. 3b) and one node is on the perpendicular side. They therefore do not require extra consideration.

**B. Polynomial-time algorithm for detecting critical regions**

As shown in Theorem 2, at most three points are needed to define a given figure. Finding the centers and the orientations (or the corners) of a certain figure requires constant time $O(1)$. There are $\binom{N}{3}$ triples of nodes and for each triple needs to consider all possible figures through this triple. If a certain node pair (or isolated node) cannot form a figure with any of the other nodes, then the algorithm for critical region

**Algorithm 1:** FINDCRITICALREGION

```
input : the network $G(N, \mathcal{L})$, figure $F$, metric $X$
output: critical region(s) $C$, metric after a failure minVal
1  C ← ∅;
2  minVal ← ∞;
3  foreach node triple $\{A, B, C\} \subseteq N$ do
4    if ($F \neq \text{ellipse and } \{A, B, C\}$ are collinear) then
5      Q ← two positions for $F$ such that one node is in the corner and all three are on the same side;
6    else
7      Q ← all positions for $F$ through $A$, $B$ and $C$;
8    foreach $q \in Q$ do
9      $G' ← G(N, \mathcal{L})$;
10     foreach $N \in N$ do
11         if $N \in q$ then $G' ← G' − N$;
12         RELAXCRITICAL($G', X, q, \text{minVal}, C$);
13    endforeach
14    endforeach
15  endforeach

Algorithm 2: RELAXCRITICAL

```

```
input : modified network $G'$, the network metric $X$, region considered $q$, current minimum minVal, critical region(s) $C$.
output: current minimum minVal, critical region(s) $C$.
1  $Y ← \text{CALCULATEMETRIC}(G', X)$;
2  if $Y < \text{minVal}$ then /* new critical region */
3    $C ← q$;
4    $\text{minVal} ← Y$;
5  else if $Y = \text{minVal}$ then /* another critical region */
6    $C ← C \cup q$;
```

detection considers a figure arbitrarily positioned through these nodes. The time-complexity of this part is $O(N^3)$. If the complexity of determining the change in the value of a metric
after the failure of a considered region is $O(C)$, then we have a complexity of $O(N^3 \cdot C)$. Checking whether a single node lies inside a positioned figure can be done in $O(1)$ time. The total complexity accumulates to $O(N^4 \cdot C)$. The algorithm is formalized in FIND CRITICAL REGION (Algorithm 1). Routine RELAX CRITICAL (Algorithm 2) is used to calculate the change in network metric $X$ in modified network $G'$ and to update the set of critical regions $C$.

IV. SIMULATION RESULTS

In this section, we conduct simulations to study how certain metrics are affected after the failure of a critical region.

For the five figures under consideration, we detect critical regions in three real-world networks: the infrastructure of the ARPANET network [14] (in Fig. 4a), which is often used as a benchmark topology, the Italian main backbone network (in Fig. 4b), and the main backbone fiber connections in Europe [15] (in Fig. 4c). We refer to these networks as: ARPANET, ITALY, and EUROPE, respectively. Through longitude and latitude information, the geographical distances between the nodes can be derived.

In ARPANET, the most critical square and circular regions are those covering the west and east coast, while the most critical stretched figures are located in central USA, because the nodes centered there link the US coasts. In ITALY, the most critical regions are positioned in the northern part of the country; however, the stretched figures touch a part of central Italy. For EUROPE, for a relatively small size of the figure, the most critical regions are situated near London.

For the same networks, we have also examined the change in two network metrics after the failure of a critical region, namely: (1) the number of disconnected pairs and (2) the average shortest path length for the five different figures. The distance control variable ($r$) is used, such that for a given $r$ the areas of different figures are the same. Fig. 5 shows that for both metrics the network is not affected equally for different figures. Generally, for the number of disconnected pairs the most critical region is more disruptive for the equilateral triangle and the stretched ones (ellipse and rectangle) than for the circle and the square (Figs. 5a, 5b and 5c).

In particular, for the average shortest path length in ARPANET (Figs. 5a and 5d), where there are distant nodes and the most dense areas in terms of nodes are not central, the most critical region is more disruptive for the equilateral triangle and the stretched figures (ellipse and rectangle), than for the circle and the square. The same holds for the number of disconnected pairs. Somewhat similar behavior is noticed for ITALY (Figs. 5b and 5e). On the other hand, for the average shortest path length in EUROPE (Figs. 5c and 5f), where there is a very dense region (United Kingdom), apart from the equilateral triangle, the square and the circle are more disruptive than the stretched figures. Indeed, when extremely large regions are considered (large $r$ in Fig. 5), the metric values for different figures become more similar as most of the nodes in the networks are affected in all cases.

Fig. 4: For the 5 figures, the critical regions (possibly multiple) as a function of the number of disconnected pairs. The area of each figure per map is the same: (a) 300020.55 km²; (b) 16895.36 km² and (c) 27906.96 km². The stretched ellipse (rectangle) has one semi-axis (side) nine times longer than the other.

V. CONCLUSION

This paper has considered the problem of finding critical network regions as a function of several two-dimensional
and equilateral triangle with sides $a$. Proof of Theorem 2

Fig. 5: The number of disconnected pairs (a), (b), (c) and the average shortest path length (d), (e), (f), with $r$ with radius $r$ and $r$ for a certain figure size. The results show that the equilateral triangle and the stretched figures might be more disruptive in terms of $x$ and $y$ results in

$$2b^2((x_A-x_m)\cos(\varphi)-(y_A-y_m)\sin(\varphi))x_j+2a^2((x_A-x_m)\sin(\varphi)+(y_A-y_m)\cos(\varphi))y_j$$

$$=(x_A-x_m)\cos(\varphi)-(y_A-y_m)\sin(\varphi)$$

where $m \in \{B,C\}$. (2) forms a system of two equations with two unknowns, treating $\varphi$ as a constant. Finding the solutions (e.g., by calculating determinants), we obtain

$$x_j = M_1(\cos(\varphi), \sin(\varphi)) y_j = N_1(\cos(\varphi), \sin(\varphi))$$

where $K_2(\cos(\varphi), \sin(\varphi))$, $M_3(\cos(\varphi), \sin(\varphi))$ and $N_3(\cos(\varphi), \sin(\varphi))$ are homogeneous polynomials in $\cos(\varphi)$ and $\sin(\varphi)$ of degree 2, 3 and 3, respectively. Using (3) in (1), for instance for node $C$, results in an equation of $\varphi$. After expanding, we end up with a homogeneous polynomial.
in $\cos(\varphi)$ and $\sin(\varphi)$ of degree 6. The last equation consists of terms $\sin^i(\varphi)\cos^{6-i}(\varphi)$ for $i = 0, 1, \ldots, 6$. We take all the terms with $i$ even on one side and all the terms with $i$ odd on the other. For the even terms: $\sin^{2l}(\varphi)\cos^{6-2l}(\varphi) = \left(\frac{1-\cos(2\varphi)}{2}\right)^l \sin^{2l}(\varphi) \cos^{6-2l}(\varphi)$ for $l = 0, 1, 2, 3$ and for the odd terms $\sin^{2l+1}(\varphi)\cos^{6-2l}(\varphi) = \frac{\sin(2\varphi)}{2} \left(\frac{1-\cos(2\varphi)}{2}\right)^l \sin^{2l+1}(\varphi) \cos^{6-2l}(\varphi)$ for $l = 0, 1, 2$. Finally, squaring both sides and using $\sin^2(2\varphi) = 1 - \cos^2(2\varphi)$ results in an equation of degree 6, solely in $\cos(2\varphi)$. The last can be solved numerically. Because of the squaring, we need to check whether the solution holds in (1). Finally, we obtain the solution via equation (3).

2) Rectangle $F_R$: There are two possible cases for a rectangle with sides lengths $a$ and $b$ (Figs. 3a and 3b):

(i) Two nodes (e.g., $A$ and $B$) lie on the same side and the third (node $C_2$) on a perpendicular side as shown in Fig. 3a.

(ii) Two nodes (e.g., $A$ and $B$) lie on parallel sides, and the third (node $C_2$) is on a perpendicular side to those sides as shown in Fig. 3b.

Now, the equations of lines through $A$, $B$, and $C$ are: $y = k_A(x-x_A) + y_A$, $y = k_A(x-x_B) + y_B$ and $y = -\frac{1}{k_A}(x-x_C) + y_C$, respectively. In this case, $k_A$ is not directly known. Corner $P$ is found as the intersection of the lines through $A$ and $C$, while corner $R$ as the intersection of the lines through $B$ and $C$. Hence, $x_P = \frac{x_Ak_A^2 + (y_C-y_A)ka + y_C}{k_A^2 + 1}$, $y_P = k_A(x_P-x_A) + y_A$ and $x_R = \frac{x_Ak_A^2 + (y_C-y_A)ka + y_C}{k_A^2 + 1}$, $y_R = k_A(x_R-x_C) + y_C$. Now, because $\mid PR\mid = \frac{a}{t}$, we end up with the equation $(x_a - x_B)^2 - t^2k_A^2 + 2(x_a - x_B)(y_B - y_A)ka + (y_B - y_A)^2 - t^2 = 0$.

Because there might be two real values of $k_A$ and two values for $t$, there are at most 4 possibilities for $k_A$. When we find a certain $k_A$, $P$ and $R$ are already determined and $Q$ and $T$ are found on appropriate distances ($a$ or $b$) from $P$ and $R$. Because, other pairs of nodes ($A$, $C_2$, $B$, or $C_2$) might lie on parallel sides, we have another 8 solutions, hence 12 (6 for the square) in total for case (ii). Finally, we have 18 cases to be checked for the rectangle (and 9 for the square).

3) Equilateral Triangle $F_T$: For an equilateral triangle, there can also be two possibilities (Figs. 3e and 3f):

(i) Two nodes (e.g., $A$ and $B$) lie on the same side in the rectangle and the third (node $C_2$) on a different side as shown in Fig. 3e. There are two possibilities for the slope of the line through $C_2$ as it closes an angle of $\frac{\pi}{3}$ or $\frac{2\pi}{3}$ with the line through $A$ and $B$. One corner $P$ is found in the intersection. Corners $Q$ and $R$ are determined on a distance $a$ from $P$. One needs to check whether $A$, $B$, and $C_2$ lie internally on the sides of $\triangle PQR$. Because we have 2 solutions in this case and there can be 3 possible pairs of nodes that lie on the same side, we have 6 potential solutions.

(ii) Nodes $A$, $B$, and $C$ lie on different sides as depicted in Fig. 3f. Assuming that the angle between the lines through $A$ and $B$ is $\frac{\pi}{3}$ and between the lines through $A$ and $C$ is $\frac{2\pi}{3}$, we obtain $k_B = \frac{k_A + \sqrt{3}}{1 - \sqrt{3}a}$, $k_C = \frac{k_A - \sqrt{3}}{1 + \sqrt{3}a}$. Now, we can find corners $P$ and $Q$ as intersections of the lines through $A$ and $B$ and through $A$ and $C$, respectively: $x_P = \frac{x_Ak_A + (y_B - y_A + \frac{\sqrt{3}}{2}a)(k_A^2 + x_B + \frac{\sqrt{3}}{2}a)}{k_A^2 + 1}$, $y_P = k_A(x_P - x_A) + y_A$ and $x_Q = \frac{x_Ak_A + (y_C - y_A + \frac{\sqrt{3}}{2}a)(k_A^2 + x_C + \frac{\sqrt{3}}{2}a)}{k_A^2 + 1}$, $y_Q = k_A(x_Q - x_A) + y_A$. The length of $PQ$ is $a$, which leads to a quadratic equation in $k_A$:

$$m^2 - a^2 k_A^2 + 2mkA + (n^2 - a^2) = 0$$

where $m = (y_B - y_C) + \frac{eA + \sqrt{3}C - 2\sqrt{3}a}{\sqrt{3}}$ and $n = (x_B - x_C) + \frac{eA + \sqrt{3}C - 2\sqrt{3}a}{\sqrt{3}}$. After $k_A$ and subsequently, $k_B$ and $k_C$ are known, corners $P$ and $Q$ are immediately found. The last corner $R$ is determined as intersection of the lines through $B$ and $C$. Finally, we need to check whether $P$, $Q$, and $R$ lie on the sides of the triangle. Because of the assumptions for the angles choice ($\frac{\pi}{3}$ or $\frac{2\pi}{3}$) and the two possible solutions in (4), we have 4 solutions in total. Based on (i) and (ii), we have 10 cases for an equilateral triangle.

REFERENCES


