Increasing Public Safety Broadband Network Resiliency Through Traffic Control

R. Rouil, W. Garey, C. Gentile, N. Golmie
National Institute of Standards and Technology, USA

P. Schwinghammer
First Responder Network Authority, USA

Abstract: Long Term Evolution (LTE)-based cellular networks are being deployed around the world to provide public safety with enhanced capabilities and access to broadband technology. In the United States, the First Responder Network Authority (FirstNet) is on the verge of deploying a nationwide network called the National Public Safety Broadband Network (NPSBN). Commercial networks typically aim at maximizing network capacity, i.e. the aggregate data rate, in order to increase revenue. In public safety networks, however, coverage, not capacity, is paramount, especially during an outage when sites are down. Through traffic control and preemption, the service level of low-priority users is reduced or denied, freeing up resources to restore coverage to high-priority users, e.g. users responding to an incident. In this study, we examine the effect of outages on network coverage and throughput. As our main contribution, we propose three traffic-control schemes that exploit variable modulation and coding, a feature that LTE enhances with respect to its 3G predecessors. The schemes differ based on the proportion of low- and high-priority users preempted. We show that indeed the network coverage can be restored significantly and we investigate the tradeoff between the three schemes. Finally, we perform sensitivity analysis to confirm the effectiveness of the schemes across a wide range of scenarios.

Keywords: network resiliency; network outage; incident; modeling

1 Introduction

The First Responder Network Authority (FirstNet) was founded in 2012 [1] to design a nationwide broadband network for first responders based on Long-Term Evolution (LTE), the driving technology for 4G wireless communications systems developed by the Third Generation Partnership Program (3GPP). The technical specifications of a public-safety-grade network have not been standardized, however the National Public Safety Telecommunications Council (NPSTC) has specified that such a network needs to provide a high level of reliability and resiliency to failures [2] in order to guarantee availability to first responders when necessary.

In recent decades, the impact of natural disasters such as hurricanes and earthquakes on telecommunications networks has been well documented. In [3], Kwasinski used field damage assessments to understand which infrastructure elements had failed, why, and how long it took to restore operation. The author examined numerous cases, including Hurricanes Katrina (2005), Gustav (2008), and Ike (2008), as well as the earthquakes/ tsunamis in Chile (2010), in Christchurch, New Zealand (2011), and in the Great Tohoku region in Japan (2011). In 2006, the Federal Communications Commission (FCC) established an independent panel to study the impact of Hurricane Katrina on telecommunications, including public-safety networks. The lessons learned are providing critical guidelines to enhance emergency communications [4], both for the radio access network (RAN) and the core network. Unfortunately, as seen recently during the flooding in Louisiana in August 2016, outages still occur and cause service disruption to thousands of users, with no exception to first responders [5].

Among the solutions used to increase RAN resiliency, hardening of cellular sites is viewed as an effective, but costly, solution; therefore, it cannot be applied to every tower. In [6], Griffith et al. developed a heuristic approach to identify high-impact sites in a given network deployment that would benefit from hardening. The use of cell-on-wheels (COWs) or cells-on-light trucks (COLTs), also called "deployables", is common practice to reestablish coverage where towers have failed [7,8], but the time required to route and setup those systems can still take several days. Kwasinski and Kwasinski in [9] demonstrate how to improve resiliency by trading off Quality-of-Service, such as data rates or delays, in order to reduce the power consumption needed by towers running on batteries or solar panels. The limitation, however, is that the work only considers power outages in which the towers, equipped with batteries, can still operate – not actual damage that would prevent operation. With the adoption of LTE and new self-optimization features introduced in LTE Advanced, novel solutions are being envisioned where the cell towers can dynamically adjust their configurations to adapt to network changes. While there are many parameters that can be adjusted, most proposed solutions only combine one or two parameters due to the rapid increase in complexity. For example, in [10], Buenestado et al. propose an algorithm focusing solely on adjusting the antenna tilts by monitoring call traces to get real time network information. The authors in [11] tune the sites based on both transmit power and antenna tilts. In both cases, the goal is to optimize coverage and capacity, suitable for normal operation but does not consider the importance of coverage during outages. In [12], Sivaraj et al. propose a dual optimization of the transmit power and beamforming pattern to minimize service degradation during an outage by offloading users to neighboring towers.

In this paper, we propose priority and preemption schemes to sustain coverage on high-priority users responding to an emergency incident for which tower outages occur. The extended coverage is provided by towers still operational, albeit removed from the incident area, by preempting service to other low-priority users. In essence, the schemes convert preempted capacity into extended coverage by exploiting advanced variable modulation and coding schemes introduced in LTE. The rest of the document is organized as follows: Section 2 describes the modeling methodology and the public-safety scenarios considered. Section 3 provides baseline results showing the impact of outages on the high-priority users when no traffic control is applied. Section 4 describes the different traffic scaling algorithms studied and their performance evaluation. Section 5 presents the results of sensitivity analyses and the last section provides concluding remarks.

2 Modeling Methodology

We describe a modeling methodology to analyze two emergency situations that fall under the theme of network resiliency: site outages and incident events. Outages may occur for many reasons. On one hand, scheduled failures may arise from maintenance issues or site relocation; in those cases, advanced planning can mitigate the resultant effects. On the other hand, there are cases where failures cannot be anticipated, including equipment lapse, theft/vandalism, administrative errors (e.g. misconfigured system parameters), and natural disasters (e.g. hurricanes, tornados, earthquakes); in the latter cases, the extent of the failure can be significant, as was the case in 2011 when 29,000 cellular base stations were damaged during an earthquake in Japan [7]. Incidents may also be caused by natural disasters – events such as building fires which trigger a high concentration of demand for wireless communications services – but do not lead to site outage. Other incidents may be hostage situations or school shootings. While incidents are typically smaller in size, both phenomena lead to greater network load: outages by reducing network capacity and incidents by boosting in situ demand.

The city of San Francisco (SF) was chosen as the analysis region for our case study, primarily due to its dense population, manifold terrain, and vulnerability to earthquakes – all challenges for network resiliency. The region is displayed in Figure 1. The analysis was performed using the Mentum Planet[1] [13] Radio-Frequency (RF) network simulator. The main simulation parameters related to channel propagation and the configuration of the base stations and user equipment are listed in Table 1. The traffic model for the first responders is derived from an incident response scenario based on the 2007 collapse of the Interstate 35 bridge in Minneapolis, Minnesota, exhibit 9[2] [14]. The scenario amounts to average downlink and uplink data rates of 18.3 kb/s and 15.3 kb/s, respectively, and represents the day-to-day demand of first responders. The distribution for the first responders, also derived from the scenario, was scaled in proportion to the population density in San Francisco and ultimately yielded a total of 574 users. The network plan was designed to achieve 95 % user coverage with reliability against large-scale (shadow) fading of 95 %, yielding 9 public-safety sites. Also displayed in Figure 1 are the public safety sites in green and the first responders in cyan.

Table 1: Network parameter configurations

Parameter	Value

[1]Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.
[2]The command-unit downlink and uplink video were removed.

Propagation model	3D raytracing model (Universal Model) [13]
Frequency bandwidth (MHz)	2x10
Standard deviation of lognormal shadow fading (dB)	7

Parameter	eNodeB Value	UE Value
Transmit power per antenna (dBm)	52.8	23
Antenna parameter	16.7 dBi boresight gain 65° azimuth / 9.7° elevation beamwidth 0° downtilt	-4.0 dBi Antenna gain
Multiple Input Multiple Output (MIMO)	2x2	1x2
Noise figure (dB)	3	9
Inter-cell interference coordination (ICIC) scheme	Static Soft Frequency Reuse (SFR)	N/A
Uplink power control	Fractional	N/A

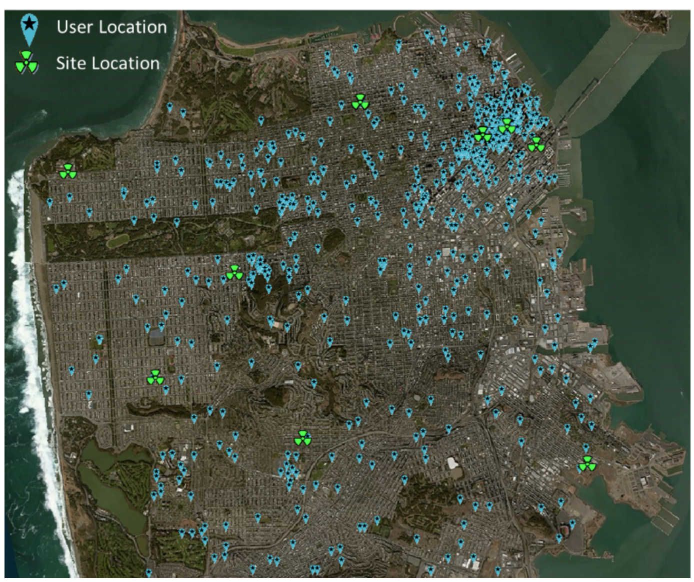
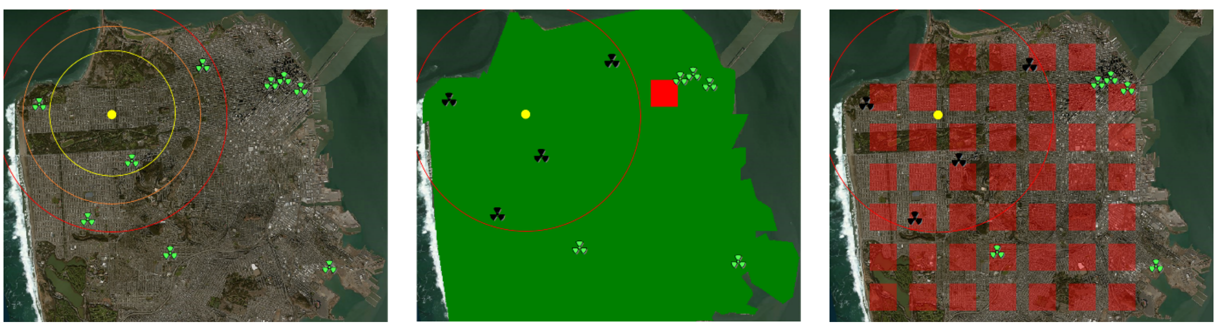


Figure 1: Public-safety network plan for San Francisco composed of 9 sites covering 574 first responders.

2.1 Resiliency analysis

Our methodology to analyze site outages is to start with the network plan and then fail (deactivate) sites that fall within some radius from the epicenter of an outage. In order to capture a wide range of potential scenarios, we vary the radius (e.g. yellow, orange, and red in Figure 2(a)) to emulate outages of different scale. Each radius has a corresponding site failure (expressed as a percentage of failed sites to total sites). The failed sites are shown in black.

During an incident, first-responder demand will rise and possibly cause congestion on the network. To model this behavior, we introduce *incident* users on the network and assign them a data rate higher than otherwise *normal* users: 32 kb/s on both the downlink and uplink. On average, there are 100 incident users (varies simulation-to-simulation) and they are contained within a 1 km x 1 km incident area. In Figure 2(b), the incident area is shown in red and the complementary area, which contains normal users only, in green. The location of the incident area was displaced per simulation according to the grid layout in Figure 2(c). The number of incident locations depends on the size of the region considered. For San Francisco, we obtained 47 incident locations.



(a) Generation of local outages (b) single instance of incident (c) location of all incidents

Figure 2: Outage and Incident modeling

When either a site outage or an incident occurs, the network becomes stressed beyond the design point; dependent on the level of stress, 95 % coverage for normal and/or incident users may no longer be achieved. When both events occur together, there is both elevated demand and reduced network capacity. As such, performance may drop precipitously, especially if they occur close to each other, i.e. when an incident area falls within an outage ring. Figure 2(b) depicts an outage occurring at the same time as an incident. In order to provide generalized results, we model both site outages and incident events simultaneously. Specifically, all performance metrics reported are averaged both over 10 randomly selected epicenters and over the 47 incident locations. The epicenters were chosen independently from the incident locations.

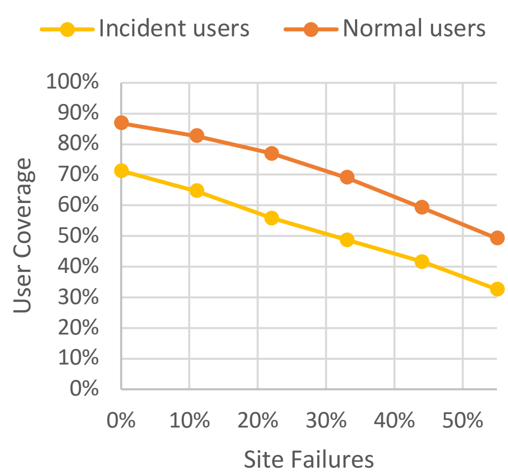
3 Network behavior during emergencies

In this section, we examine the impact of site outages and incidents on the first responders in the region. Impact can be gauged through two performance metrics we have identified:

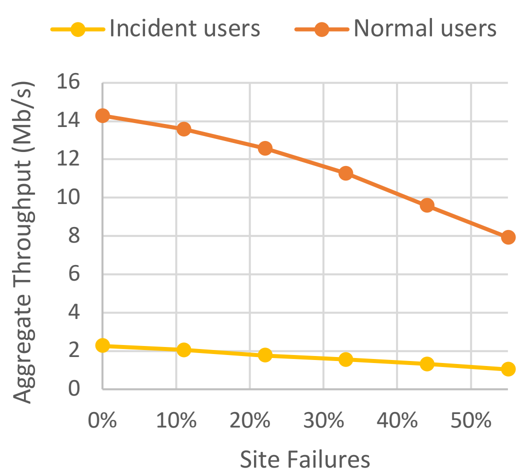
- *User coverage*: ratio of users served to total number of users deployed (expressed in percentage). A user is served if its throughput demand can be met by the network; otherwise its throughput is zero.

- *Aggregate throughput*: sum of individual user throughput over users deployed[3]. When the network is saturated, this quantity corresponds to the network capacity.

Figure 3 displays the performance metrics versus site failure percentage for both normal and incident users. Note that the coverage and throughput curves are proportional to each other; this is because here coverage translates directly to throughput through the user data rate. Since priority and preemption have not been implemented, both user types are impacted. A key observation is that only 87 % of the normal users are covered when there is no outage (0 % site failure); in other words, the 95 % coverage criteria cannot be achieved. This indicates that in some areas the network cannot handle the surge of traffic generated by the incident. Indeed, the incident users are covered at a yet lower percentage (70 %) because their demand is higher. As expected, the curves drop off as the outage intensifies due to ever reducing capacity.



(a) User coverage



(b) Aggregate throughput

Figure 3: Impact of outages on normal and incident users without traffic control

[3]This is equivalent to the sum of individual user throughput over users served only since users not served have zero throughput.

To further illustrate the impact of outage on incident user coverage, the three subfigures (a-c) in Figure 4 synthesize the coverage maps of the 47 individual incident areas for a given site outage configuration. The active / failed sites for a configuration are displayed in each subfigure as is the degree of coverage for each incident area color-coded against the legend. As expected, coverage is worst in the vicinity of the failed sites. Less expected is the drop in coverage surrounding the active sites as well. This is explained by the fact that the active sites, in the wake of an outage, have to handle not only the incident users but also the normal users affected by the outage. This behavior is described in detail in the next subsection.

(a) outage: 11 % (b) outage: 33 % (c) outage: 66 %

Figure 4: Maps showing incident user coverage for select outages

3.1 Understanding network dynamics during emergencies

In LTE terminology, the mobile device is referred to as a User Equipment (UE) and a base station as an Evolved Node B (eNodeB). Both will have a predetermined number of radio resources given by the allocated bandwidth of operation. Signal-to-Interference-and-Noise Ratio (SINR) is an indicator of the signal quality from an eNodeB: typically, the closer the UE is to the eNodeB, the greater the signal strength is and the lower the interference from other stations, resulting in higher SINR. LTE differs from its 3G predecessors in many ways. Relevant to our work here is that it features variable, not fixed, modulation and coding schemes (MCSs). In practice, each MCS has an associated data rate per radio resource, otherwise known as *spectral efficiency*. If the SINR is strong, the signal is more robust to decoding errors and data, in turn, can be transmitted at a higher rate. A direct consequence is that a user farther from the eNodeB will require more radio resources than a user close to the eNodeB for the same data rate. This is illustrated in the following example.

Figure 5 shows a network with incident and normal users served by base stations A and B. Ideally, the individual coverages of the stations will overlap in order to provide seamless service (handoff) to users at cell edge. For convenience, a simplified representation of the SINR and spectral efficiency as a function of user distance from the stations are also shown. To maximize spectral efficiency (minimize radio resources), the users are assigned to cells A and B accordingly.

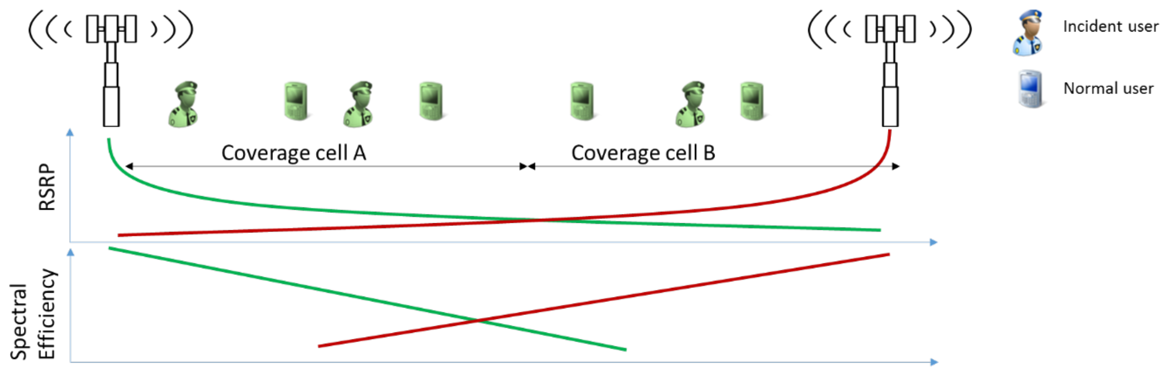


Figure 5: User coverage without outage

Besides for handoff, the degree of cell overlap will also determine whether a user can switch to a neighboring site in case of outage. For instance, consider station A in Figure 6. As a result of outage, users in cell A will attempt to reconnect to station B: some will be able to detect station B but will be unable to connect due to their positions at cell edge, at which their spectral efficiency requires more resources than what are available; yet other users will not detect the station at all because they are too far away (all disconnected users from cell A are shown in red). The remaining users will be able to reconnect and sustain connectivity, however, by doing so will take away resources from UEs that were originally served by station B. This will cause some of them to disconnect or have their service degraded (shown in orange). This demonstrates how users in the vicinity of the active cells in Figure 4 lose coverage. Ultimately, users served by station B will be selected based on admission control and scheduling protocols implemented by the LTE network. Through the same mechanisms, the LTE network has the ability to preempt, or reduce service to, low-priority (normal) users in the favor of high-priority (incident) users. This is the topic of the next section.

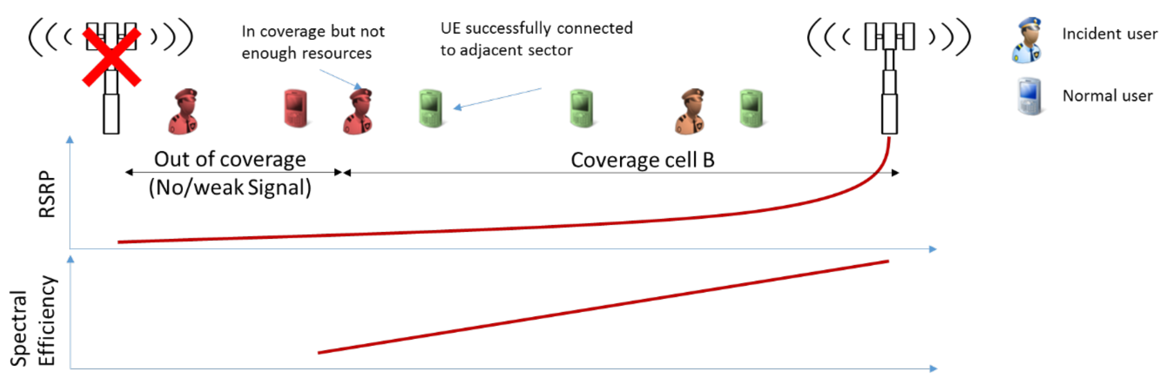


Figure 6: User coverage with outage and no traffic control

4 Traffic-control Schemes

As illustrated in the previous section, outages can be detrimental to user coverage. Priority and preemption schemes can be implemented to restore coverage to incident users during an emergency by reallocating available capacity away from normal users [15]. To date, these schemes are limited in public (commercial) networks. More specifically, "ruthless preemption" used by public safety in Land Mobile Radio (LMR) systems, is not provisioned in commercial networks; in fact, even emergency alert messages are prohibited from preempting an active voice or data session [16]. At best, priority queueing is employed for services such as Wireless Priority Service (WPS) that are designed for national security and emergency preparedness (NS/EP) users [17].

Even if the network is able to implement priority and preemption, the challenge of striking a balance between capacity and coverage remains. In commercial networks, the end goal is to maximize capacity (aggregate throughput), hence eNodeBs will prefer to serve UEs that have the best spectral efficiency (those closest to the eNodeB); this, however, will result in "shrinkage" of cell coverage. Conversely, in public safety networks the goal is to cover all first responders – incident users with higher priority – and so the eNodeBs must serve UEs further away as well; this, however, will result in reduced capacity since those UEs have lower spectral efficiency.

The principle of trading coverage for capacity, referred to as "cell breathing", is studied in [18,19]. Specifically, in [18] Jaber et al. show the effect of loading on the cell range and its impact on network dimensioning. In [19], Yang et al. propose a load balancing algorithm to move users from a congested cell to neighboring cells that may be underutilized. However, those studies focus on controlling the sector load and do not address the issue of maintaining coverage for a specific set of users during an outage.

In this section, we propose traffic-control schemes that operate on the principle of cell breathing. Specifically, normal users are preempted – either forced to operate at reduced data rates or denied service altogether – in order to free up resources for incident users whose coverage has been affected by an outage. This mechanism is illustrated in Figure 7. Two normal users in cell B are denied service (shown in red) in order to serve an incident user (shown in green) previously in cell A. Because the incident user lies farther away than the normal users, it requires resources from two normal users. Effectively, the capacity of the normal users is traded to extend coverage to the incident user. An additional benefit of traffic control is lower intercell interference resulting from relaxed network load. This supports an increase in SINR, thereby improving the spectral efficiency of the remaining users, a phenomenon called "cell elasticity" [18].

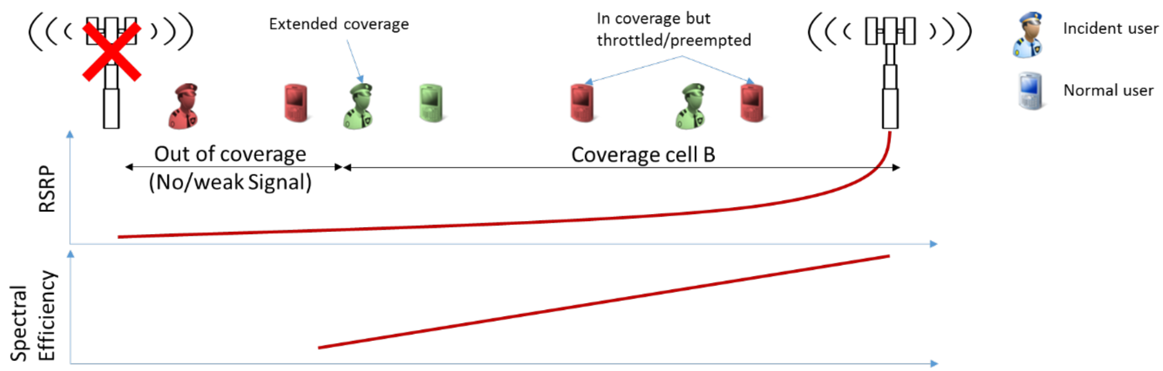


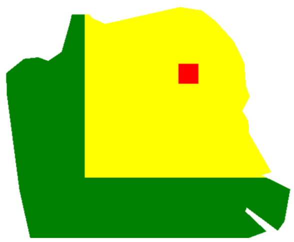
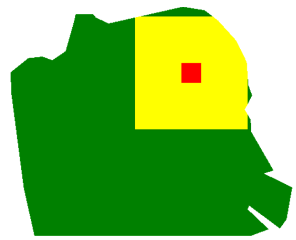
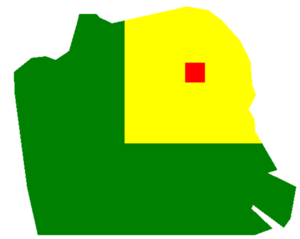
Figure 7: User coverage with outage and traffic control

In the sequel, details of the three schemes – known as Surrounding Traffic Scaling (STS), Uniform Traffic Scaling (UTS), and Incident-to-Surrounding Traffic Scaling (ISTS) – are described. The intention of our work is to evaluate the schemes to understand the factors impacting their performance. While the underlying objective to sustain coverage on incident users is common to all three, they differ in the priority assigned to incident and normal users.

4.1 Surrounding Traffic Scaling (STS)

The objective of the STS algorithm is to scale (preempt through denial-of-service) as many normal users as necessary in order to deliver 95 % coverage to incident users during an emergency. The sectors serving the incident users will be the most stressed; as such, normal users in the periphery of the incident are targeted in order to free up resources of those sectors to allocate towards incident users. Accordingly, the presupposed shape of the surrounding area, as the incident area, is also square and its dimensions are found through a binary search. The objective of the search is to determine the square for which incident user coverage is maintained at 95 % while all users in the periphery are denied service. Table 2 illustrates three iterations of the binary search: the first iteration provides incident coverage exceeding 95 %, meaning that a superfluous number of users have been scaled; the next iteration swings to the other extreme, for which users need yet to be scaled; finally, the last iteration provides the desired solution. An illustration of the STS algorithm, in comparison to the Base scheme in Figure 8(a) for which no users are scaled, appears in Figure 8(b).

Table 2: Example for computing the surrounding area to apply traffic control

Incident Normal scaled Normal	ITERATION 1 		ITERATION 2 		ITERATION 3 	
User type	Number	Coverage (%)	Number	Coverage (%)	Number	Coverage (%)
Incident	100	100	100	73	100	95
Normal scaled	442		346		430	
Normal	105		201		117	

4.2 Uniform Traffic Scaling (UTS)

While effective, the STS algorithm accounts for coverage of the incident users only, placing the burden of preemption on normal users alone. A more balanced approach is to scale incident and normal users proportionally. That is what the UTS algorithm does. Specifically, the same percentage of incident users and peripheral normal users are scaled such that coverage for incident users is restored to 95 %. Figure 8(c) illustrates a preemption of 50 % compared to Figure 8(a).

4.3 Incident-to-Surrounding Traffic Scaling (ISTS)

The ability to trade capacity will be limited by the total number of active sites in the network and so the objective to cover 95 % of incident users cannot always be met. This will typically occur only for outages with very high site failure. When this does occur, even scaling all normal users will not restore incident coverage to 95 %. It will therefore be necessary to scale some of the incident users as well such the coverage of the remaining incident users, at least, can be restored to 95 %. The preempted incident users are chosen randomly in our example; in a real environment, however, the selection would be based on priorities assigned to the users factoring in user type, application type, and/or location [15]. This is what we refer to as the ISTS algorithm and it is invoked only if the STS fails. Figure 8(d) illustrates preempting 25 % of the incident users.

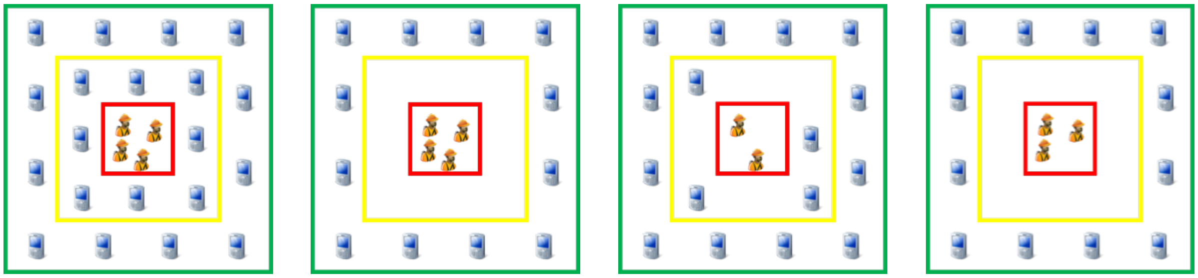


Figure 8: Illustration of different traffic scaling algorithms used

4.4 Algorithm Analysis

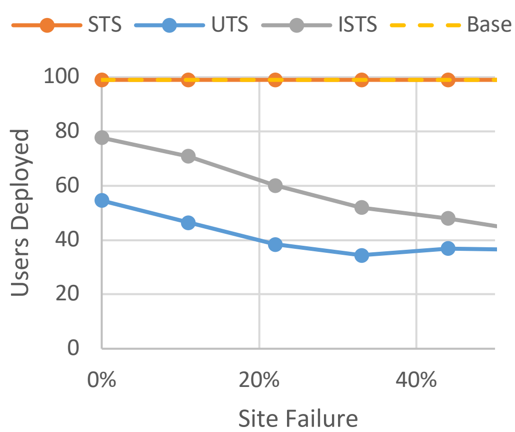
We compare the STS, UTS, and ISTS schemes to the Base scheme. To do this effectively, it is necessary to introduce a final performance metric:

- *Users deployed:* number of users that will attempt to connect to the network. Essentially, this is the number of users that are not scaled.

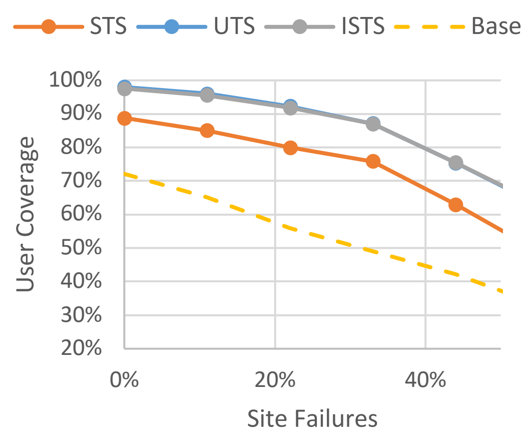
Figure 9(a) shows the incident users deployed versus the site failure percentage. True to form, the users deployed remains constant for both the Base and the STS. Conversely, in the other two schemes, the incident users are scaled back as well, albeit to a lesser degree in the ISTS (only when necessary) than the UTS. As explained in Section 3, a drop in incident-user coverage is witnessed in Figure 9(b) for the Base. By scaling normal users, the STS is capable of restoring a significant amount of that coverage, from 49 % to 76 % at 33 % site failure, though never above 95 %. The scaling results in less aggregate load on the network and in turn less interference on the remaining users. With lower interference, users can operate at higher SINR and so higher spectral efficiency. And, since the number of incident users remains constant with the STS, the aggregate throughput actually rises with respect to the Base, as shown in Figure 9(c). The ISTS and UTS push that coverage yet higher, maintaining 95 % coverage for up to 10 % site failure and providing 87 % coverage at 33 % site failure, because the incident users are also scaled back. The distinction is that the ISTS is able to serve more incident users compared to the UTS – 52 vs. 34 at 33 % failure – to the detriment of normal users – 183 vs. 267. It is important to understand that traffic scaling improves the spectral efficiency not just for the STS, but for the ISTS and the UTS as well. Hence for them, any drop in aggregate throughput with respect to the Base is by virtue of the reduction in the incident users deployed; and because of this reduction, the throughput is no longer proportional to the coverage as for the Base and the STS.

The same performance metrics for the normal users instead are shown in Figure 9(d-f). When comparing Figure 9(d) to Figure 9(a), we can see that, with traffic control, as normal users deployed decrease, incident users deployed increase. The normal user coverage shown in Figure 9(e) also increases for all traffic scaling algorithms. This is expected because reducing the network load by scaling normal users affects all users on the network. Finally, by examining Figure 9(f) and Figure 9(b), we can clearly see how the scaling algorithms, in reference to the baseline scenario, trade capacity of normal users for coverage of incident users.

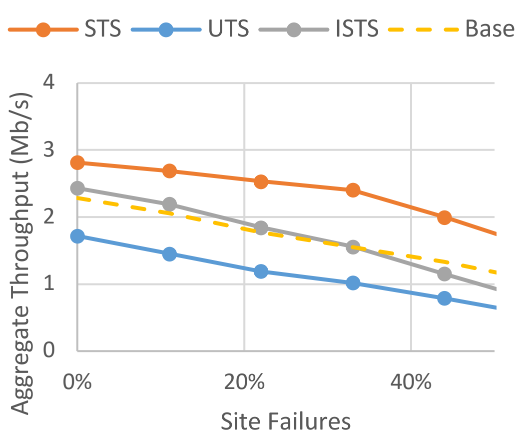
The analysis highlights the flexibility in applying different scaling to incident and normal users in order to achieve a certain network objective: at one extreme, the STS admits the most incidents users on the network to the detriment of incident-user coverage while at the other extreme the UTS achieves the converse objective; in the middle, the ISTS achieves a compromise between the two.



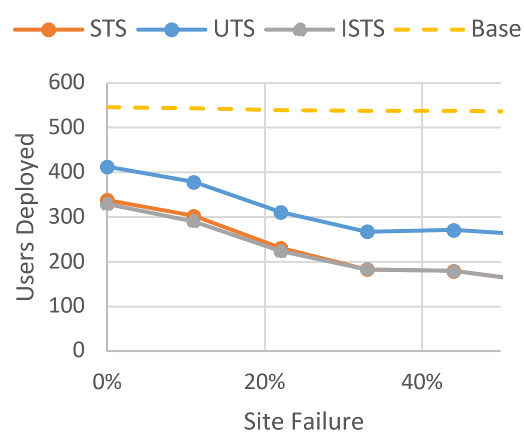
(a) Incident users deployed



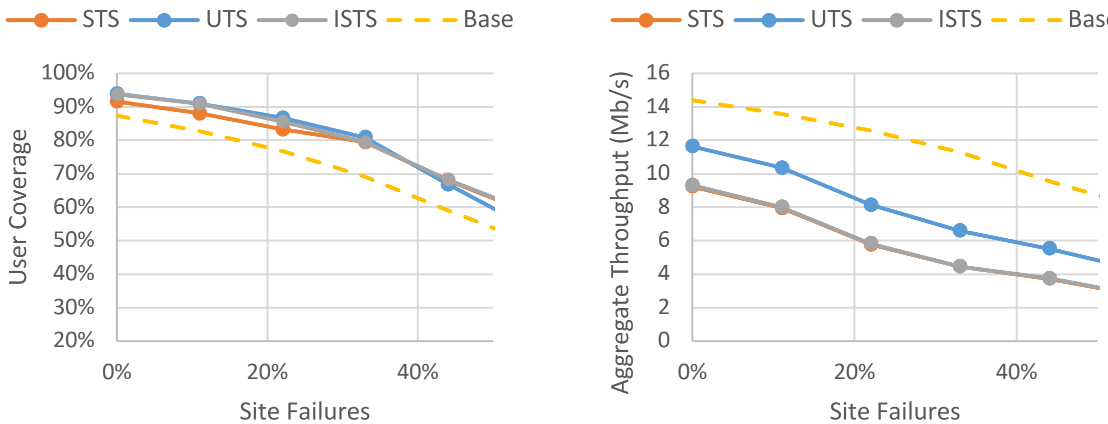
(b) Incident users' coverage



(c) Incident user throughput



(d) Normal users deployed



(e) Normal users' coverage (f) Normal user throughput

Figure 9: Incident users deployed during outage

5 Sensitivity Analysis

The previous section shows encouraging results for sustaining coverage on first responders during an emergency through traffic control. To substantiate the results further, in this section we conduct sensitivity analysis to evaluate the impact of certain simulation parameters on the performance of the traffic-control schemes. In the analysis, we vary a single parameter from its nominal value used this far while keeping all others constant. The four parameters for variation are user density, user demand, outage distribution, and analysis region. To limit the number of performance curves presented, we concentrate on the ISTS only.

5.1 Impact of user density

Although enhancing infrastructure is costly, operators do update their networks regularly. This is justified by the need for higher capacity, i.e. delivering faster data rates or serving a greater user population. In reference to the latter, the first parameter we consider in in our sensitivity analysis is the population density of normal users. It is multiplied by a factor of 4x and 8x while maintaining the number of incident users fixed at 100. Because users must still be covered at 95 %, more sites will be required to support the additional users, hence each factor will have a companion design. Figure 10 shows the region map overlaid with the user distribution and site locations for the baseline and the two factors considered.

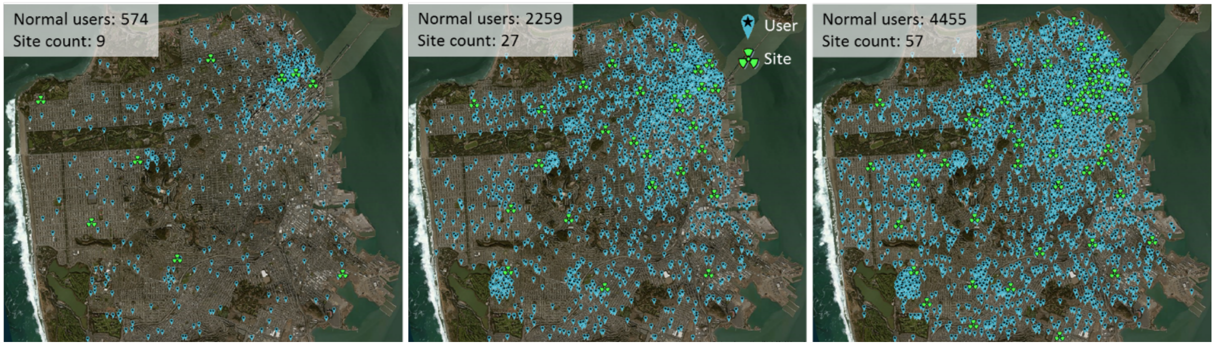
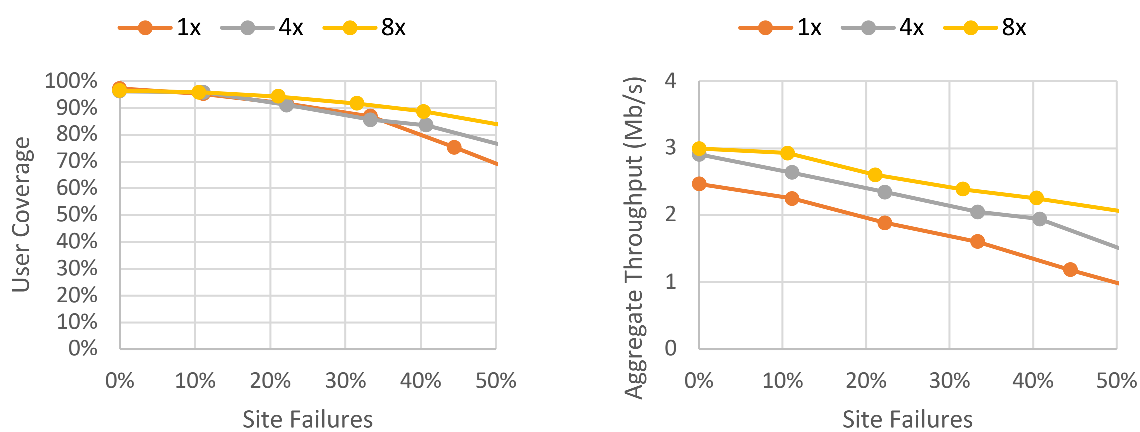


Figure 10: Network deployments supporting 1x, 4x, and 8x more normal users

We can observe in Figure 11(a) that network plans designed to handle more traffic allow for better incident-user coverage, especially for larger outages. At 40 % site outage, 1x plan provides 80 % coverage while the coverage for 4x and 8x increases to 84 % and 89 %, respectively. The performance gains are more remarkable considering that a 40 % site outage represents only 4 sites with 1x compared to 11 and 23 sites for 4x and 8x. This is because denser networks make for greater overlap between sites, improving the chances for a UE to find an alternative eNodeB to connect to in case the current one fails. Greater throughput for incident users is also witnessed with greater traffic in Figure 11(b). While adding more sites improves network resiliency, it is costly, and the number of first responders is usually dwarfed by commercial users. One approach would be to share extra capacity with non-public safety users, albeit assigned a low priority in case of an emergency.

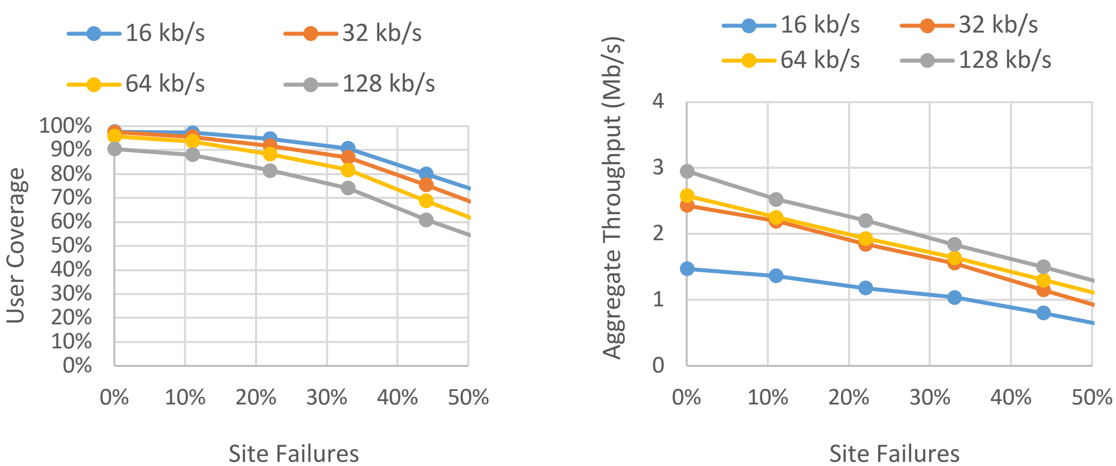


(a) Incident user coverage (b) Aggregate throughput

Figure 11: Impact of site plan on incident users with ISTS algorithm

5.2 Impact of user demand

The purpose of traffic control in our application is to trade normal-user capacity for incident-user coverage. The amount of capacity required for trade will depend on the incident-user demand. In our sensitivity analysis, we vary demand through the user data rate as 16 kb/s, 32 kb/s (nominal), 64 kb/s, and 128 kb/s for the ISTS. Its impact on coverage is displayed in Figure 12(a): with no outage, the ISTS is able to maintain coverage for all data rates up to 64 kb/s, but then coverage drops to 90 % at 128 kb/s; the trend of lower coverage vs. higher date rate is consistent for all site failures investigated. The impact on throughput in Figure 12(b) follows: with 0 % site failure, the incident user throughput increases from 1.5 Mb/s to 3 Mb/s when boosting the data rate from 16 kb/s to 128 kb/s; hence, even though more users need to be scaled, the throughput per user is higher and so the combined effect is more aggregate throughput. As outages intensify, performance degrades further. Conversely, performance improves when the data rate is reduced. This suggests that the other preemption mechanism mentioned in Section 4 – reducing data rates – is a viable alternative to denial-of-service through scaling. In practice, the former can be realized by tapering application rates (e.g. videos) or limiting the types of applications available to the users (e.g. voice only). As far as the network is concerned, the degree of scaling vs. data-rate reduction will not matter and is simply an implementation preference.

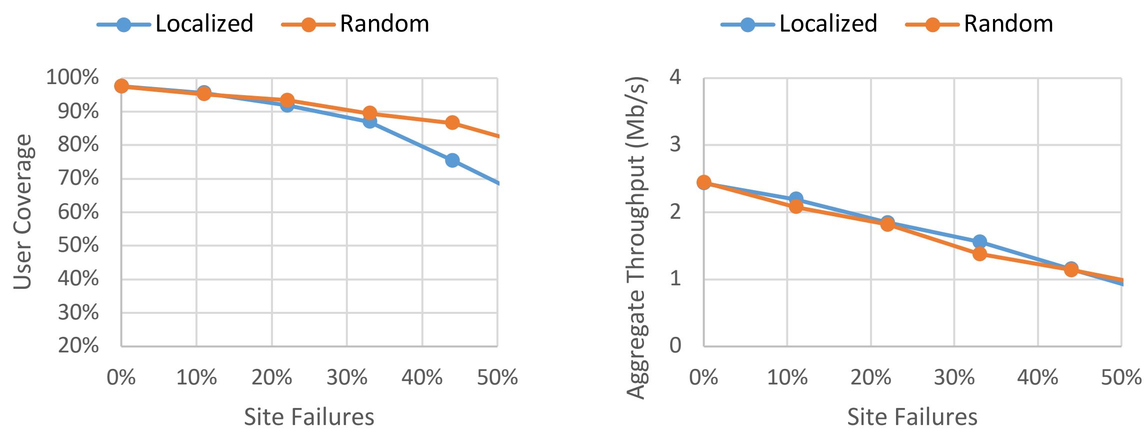


(a) User coverage (b) Aggregate throughput

Figure 12: Impact of incident-user demand on incident users on ISTS algorithm

5.3 Impact of outage distribution

Thus far, we have considered only outages surrounding the epicenter of a disaster such as a tornado, for which damages are localized to the area. Due to terrain, specific location and hardening of sites, and type of event, it is possible that the distribution of failed sites is instead sparse. In light of this, rather than a localized distribution, in our perturbation analysis we also consider uniform random distribution for outages throughout the region. Figure 13 compares results between the two. We observe that for random outages, the ISTS is more effective in restoring user coverage while providing the same amount of capacity to the incident users. The performance difference is more noticeable at higher outages: at 10 % site failure, the coverage is 95 % for both localized and random while it decreases to 75 % to 86 %, respectively, at 44 % failure. The difference is due to the fact that for random outages, the remaining sites are distributed uniformly and so are more likely to provide some level of coverage throughout the entire geographical region.



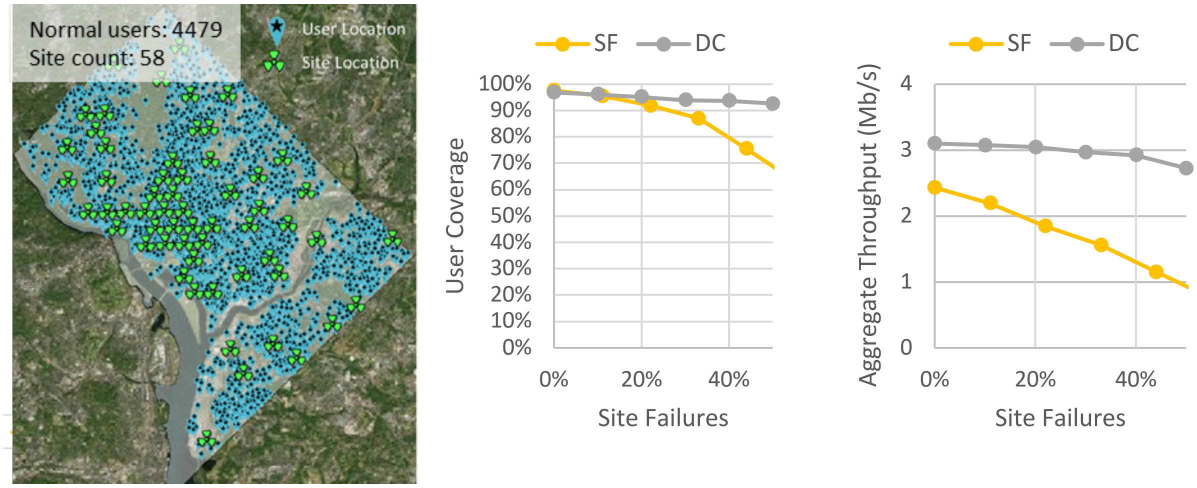
(a) User coverage (b) Aggregate throughput

Figure 13: Impact of outage distribution on incident users with ISTS algorithm

5.4 Impact of analysis region

The final variable considered in our sensitivity analysis is the analysis region. The characteristics specific to a region can impact network performance significantly. In particular, terrain and clutter will affect propagation and in turn throughput and coverage via the SINR. Besides its importance to public safety as the nation's capital, we selected the District of Columbia (DC) as an alternative region for analysis due to the flatness of its terrain in comparison to the hills of San Francisco; moreover, in stark contrast to the tall skyline in SF, especially in the business district, no skyscrapers are permitted within the city proper. Finally, the population density in DC is about half that of SF. The site plan in DC is shown in Figure 14(a) and the resultant analyses in Figure 14(b) and Figure 14(c). What is clear from the results is that the site plan in DC is more robust to site failure: in DC, the ISTS scheme is able to sustain incident-user coverage above 95 % for all 100 users up to 20 % failure. The situation is very different in SF: when exposed to an incident, minimum user coverage drops below 92 % at 20 % failure. Naturally, the performance degrades with increasing outage, but the drop-off is much steeper in SF than DC, both in terms of the coverage and throughput.

The disparity in performance has mostly to do with the different terrain and clutter between the two cities: because the landscape is flat in DC, more favorable propagation conditions exist due to the prevalence of line-of-sight throughput. Hence when an outage occurs, adjacent sites farther out can still provide coverage because the obstructions are less severe. As a further confirmation of this, it is interesting to point out that both the number of users deployed and the number of sites planned in DC (4479 users and 58 sites) are comparable to the 8x plan in SF shown in Section 5.1 (4455 users and 57 sites). The coverage area, however, is much more extensive in DC (177 km^2) than SF (11 km^2) because of the better propagation conditions.



(a) Public-safety network plan for DC (b) User coverage (b) Aggregate throughput

Figure 14: Impact of area on incident users with ISTS algorithm

6 Conclusion

In this report, we demonstrate the impact that network outages can have on the coverage of LTE broadband networks in application to the public-safety network being deployed by FirstNet. To mitigate the impact, we propose three traffic-control schemes that exploit the cell-breathing features of LTE, i.e. trading capacity of low-priority users to extend coverage to high-priority users (first responders assigned to an incident). The results evidence that network resiliency, quantified by the coverage of high-priority users, can be enhanced dramatically by scaling low-priority traffic. Depending on the severity of the outage and incident at hand, it is sometimes necessary to also scale high-priority traffic as well in order to ensure reliable coverage. We also performed sensitivity analysis to substantiate the validity of the results across a wide range of scenarios. Specifically, we varied the site density, the application data rates, the outage distribution, and the analysis region. Unsurprisingly, we observed that network resiliency improves when adding capacity to the network in terms of sites. Less trivial is the idea that building a public-safety network with sufficient capacity to support commercial users as well can provide greater resiliency than a smaller network dedicated to public-safety users alone. Future work includes the development of practical solutions and protocols to determine the area and amount of users to scale in real-time in the field.

References

[1] P. Law, Middle Class Tax Relief and Job Creation Act of 2012, 2012.

[2] National Public Safety Telecommunications Council, Defining Public Safety Grade Systems and Facilities, 2014. http://www.npstc.org/download.jsp?tableId=37&column=217&id=3066&file=Public_Safety_Grade_Report_140522.pdf.

[3] A. Kwasinski, Effects of notable natural disasters from 2005 to 2011 on telecommunications infrastructure: Lessons from on-site damage assessments, INTELEC, Int. Telecommun. Energy Conf. (2011). doi:10.1109/INTLEC.2011.6099777.

[4] N.J. Victory, Independent Panel Reviewing the Impact of Hurricane Katrina on Communications Networks - Report and Recommendations to the Federal Communications Commission, … Commun. Comm. Washington, DC. (2006) 53. http://transition.fcc.gov/pshs/docs/advisory/hkip/karrp.pdf.

[5] Inside Towers, AT&T Outages Add To Louisiana Flooding Woes, (2016). https://insidetowers.com/cell-tower-news-att-outages-add-to-louisiana-flooding-woes (accessed September 14, 2016).

[6] D. Griffith, R. Rouil, A. Izquierdo, N. Golmie, Measuring the Resiliency of Cellular Base Station Deployments, (2015) 1643–1648. doi:10.1109/WCNC.2015.7127711.

[7] T. Sakano, Z. Fadlullah, T. Ngo, H. Nishiyama, M. Nakazawa, F. Adachi, N. Kato, A. Takahara, T. Kumagai, H. Kasahara, S. Kurihara, Disaster-resilient networking: A new vision based on movable and deployable resource units, IEEE Netw. 27 (2013) 40–46. doi:10.1109/MNET.2013.6574664.

[8] M. Casoni, C.A. Grazia, M. Klapez, N. Patriciello, A. Amditis, E. Sdongos, Integration of satellite and LTE for disaster recovery, IEEE Commun. Mag. 53 (2015) 47–53. doi:10.1109/MCOM.2015.7060481.

[9] A. Kwasinski, A. Kwasinski, Tradeoff between Quality-of-Service and resiliency: A mathematical framework applied to LTE networks, 2016 IEEE Int. Conf. Commun. Work. ICC 2016. (2016) 348–353. doi:10.1109/ICCW.2016.7503812.

[10] V. Buenestado, M. Toril, S. Luna-Ramirez, J.M. Ruiz-Aviles, A. Mendo, Self-Tuning of remote electrical tilts based on call traces for coverage and capacity optimization in LTE, IEEE Trans. Veh. Technol. PP (2016) 1–13. doi:10.1109/TVT.2016.2605380.

[11] A. Engels, M. Reyer, X. Xu, R. Mathar, J. Zhang, H. Zhuang, Autonomous self-optimization of coverage and capacity in LTE cellular networks, IEEE Trans. Veh. Technol. 62 (2013) 1989–2004. doi:10.1109/TVT.2013.2256441.

[12] R. Sivaraj, I. Broustis, N.K. Shankaranarayanan, V. Aggarwal, P. Mohapatra, Mitigating Macro-Cell Outage in LTE-Advanced Deployments, (2015) 1284–1292.

[13] InfoVista, Mentum Planet: RF Network Planning & Optimization by InfoVista, (2017).

[14] T. Peha, Jon M.; Johnston, Walter; Amodio, Pat; Peters, The Public Safety Nationwide Interoperable Broadband Network: A New Model for Capacity, Performance and Cost, Fed. Commun. Comm. Washington, DC. (2010).

[15] National Public Safety Telecommunications Council, Priority and Quality of Service in the Nationwide Public Safety Broadband Network, (2015).

[16] United States Government Publishing Office, CODE OF FEDERAL REGULATIONS, 2017.

https://ecfr.io/Title-47/pt47.1.10#se47.1.10_1510.

[17] Nyquetek Inc., Wireless Priority Service for National Security / Emergency Preparedness: Algorithms for Public Use Reservation and Network Performance, (2002). http://wireless.fcc.gov/releases/da051650PublicUse.pdf.

[18] M. Jaber, Z. Dawy, N. Akl, E. Yaacoub, Tutorial on LTE/LTE-A Cellular Network Dimensioning Using Iterative Statistical Analysis, IEEE Commun. Surv. Tutorials. 18 (2016) 1355–1383. doi:10.1109/COMST.2015.2513440.

[19] S. Yang, W. Zhang, X. Zhao, Virtual cell-breathing based load balancing in downlink LTE-A self-optimizing networks, 2012 Int. Conf. Wirel. Commun. Signal Process. WCSP 2012. (2012). doi:10.1109/WCSP.2012.6543014.