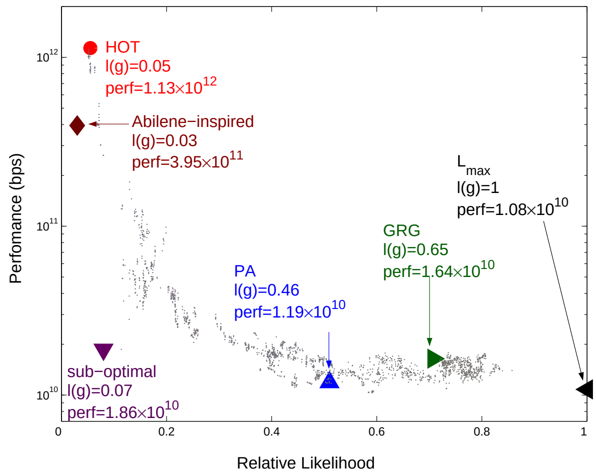

Figure 8: Performance vs. Likelihood for each topology, plus other networks having the same node degree distribution obtained by pairwise random rewiring of links.

the real Internet, aggregates traffic and disperses it across multiple high-bandwidth routers. We calculate the distribution of end user bandwidths and router utilization when each network achieves its best performance. Figure 7 (a) shows that the HOT network can support users with a wide range of bandwidth requirements, however the PA and GRG models cannot. Figure 7(d) shows that routers achieve high utilization in the HOT network, whereas, when the high degree "hubs" saturate in the PA and GRG networks, all the other routers are left under-utilized (Figure 7(b)(c)). The networks generated by these two degree-based probabilistic methods are essentially the same in terms of their performance.

Performance vs. Likelihood. A striking contrast is observed by simultaneously plotting performance versus likelihood for all five models in Figure 8. The HOT network has high performance and low likelihood while the PA and GRG networks have high likelihood but low performance. The interpretation of this picture is that a careful design process explicitly incorporating technological constraints can yield high-performance topologies, but these are extremely rare from a probabilistic graph point of view. In contrast, equivalent power-law degree distribution networks constructed by generic degree-based probabilistic constructions result in more likely, but poor-performing topologies. The "most likely" L_{max} network (also plotted in Figure 8) has poor performance.

This viewpoint is augmented if one considers the process of pairwise random degree-preserving rewiring as a means to explore the space of graphs having the same overall degree distribution. In Figure 8, each point represents a different network obtained by random rewiring. Despite the fact that all of these graphs have the same overall degree distribution, we observe that a large number of these networks have relatively high likelihood and low performance. All of these graphs, including the PA and GRG networks, are consistent with the so-called "scale-free" models in the sense that they contain highly connected central hubs. The fact that there are very few high performance graphs in this space is an indication that it would be "hard" to find a relatively good design using random rewiring. We also notice that low likelihood itself does not guarantee a high performance network, as the network in Figure 6(f) shows that it is possible to identify probabilistically rare and poorly performing networks. However, based on current evidence, it does appear to be the case that it is impossible using existing technology to construct a network that is both high performance and high likelihood.

5.2 A Second Example

Figure 6 shows that graphs having the same node degree distribution can be very different in their structure, particularly when it comes to the engineering details. What is also true is that the same core network design can support many different end-user bandwidth distributions and that by and large, the variability in end-user bandwidth demands determines the variability of the node degrees in the resulting network. To illustrate, consider the simple example presented in Figure 9, where the same network core supports different types of variability in end user bandwidths at the edge (and thus yields different overall node degree distributions). The network in Figure 9(a) provides uniformly high bandwidth to end users; the network in Figure 9(b) supports end user bandwidth demands that are highly variable; and the network in Figure 9(c) provides uniformly low bandwidth to end users. Thus, from an engineering perspective, not only is there not necessarily any implied relationship between a network degree distribution and its core structure, there is also no implied relationship between a network's core structure and its overall degree distribution.

6. DISCUSSION

The examples discussed in this paper provide new insight into the space of all possible graphs that are of a certain size and are constrained by common macroscopic statistics, such as a given (power law) node degree distribution. On the one hand, when viewed in terms of the (relative) likelihood metric, we observe a dense region that avoids the extreme ends of the likelihood axis and is populated by graphs resulting from random generation processes, such as PA and GRG. Although it is possible to point out details that are specific to each of these "generic" or "likely" configurations, when viewed under the lens provided by the majority of the currently considered macroscopic statistics, they all look very similar and are difficult to discern. Their network cores contain high connectivity hubs that provide a relatively easy way to generate the desired power law degree distribution. Given this insight, it is not surprising that theorists who consider probabilistic methods to generate graphs with power-law node degree distributions and rely on statistical descriptions of global graph properties "discover" structures that are hallmarks of the degree-based models.

However, the story changes drastically when we consider network performance as a second dimension and represent the graphs as points in the likelihood-performance plane. The "generic" or "likely" graphs that make up much of the total configuration space have such bad performance as to make it completely unrealistic that they could reasonably represent a highly engineered system like an ISP or the Internet as a whole. In contrast, we observe that even simple heuristically designed and optimized models that reconcile the tradeoffs between link costs, router constraints, and user traffic demand result in configurations that have high performance and efficiency. At the same time, these designs are highly "non-generic" and "extremely unlikely" to be obtained by any random graph generation method. However, they are also "fragile" in the sense that even a small amount of random rewiring destroys their highly designed features and results in poor performance and loss in efficiency. Clearly, this is not surprising—one should not expect to be able to randomly rewire the Internet's router-level connectivity graph and maintain a high performance network!

One important feature of network design that has not been addressed here is *robustness* of the network to the failure of nodes or links. Although previous discussions of robustness have featured prominently in the literature [4, 42], we have chosen to focus on the story related to performance and likelihood, which we believe