The Intrinsic Scale Of Networks Is Small

Malik Magdon-Ismail
Computer Science Department
Rensselaer Polytechnic Institute
Troy, NY, USA
magdon@cs.rpi.edu

Kshiteesh Hegde
Computer Science Department
Rensselaer Polytechnic Institute
Troy, NY, USA
hegdek2@rpi.edu

Abstract—We define the intrinsic scale at which a network begins to reveal its identity as the scale at which subgraphs in the network (created by a random walk) are distinguishable from similar sized subgraphs in a perturbed copy of the network. We conduct an extensive study of intrinsic scale for several networks, ranging from structured (e.g. road networks) to ad-hoc and unstructured (e.g. crowd sourced information networks), to biological. We find: (a) The intrinsic scale is surprisingly small (7-20 vertices), even though the networks are many orders of magnitude larger. (b) The intrinsic scale quantifies "structure" in a network – networks which are explicitly constructed for specific tasks have smaller intrinsic scale. (c) The structure at different scales can be fragile (easy to disrupt) or robust.

Index Terms—structure, signature, deep learning, convolutional neural networks

I. INTRODUCTION

Large networks are ubiquitous, either explicitly (e.g. the Facebook social network) or implicitly (e.g. the DBLP citation data induces a network of researchers; the Amazon purchase data induces a product network). Significant effort has been spent quantifying a network's topological structure. Seidman [21] computes network cohesion using minimum vertex cuts. Reagans et al [19] view network structure as facilitating knowledge transfer and argue that social ties, cohesion and network range play important roles. Olbrich et al [18] use exponential families to estimate degree distribution, clustering and assortativity coefficients, and subgraph densities. Clustering the vertices based on the topology is a powerful tool for uncovering structure. Newman in [16], [17] developed a popular approach to non-overlapping clustering, which optimizes a modularity objective that (globally) quantifies the quality of the entire collection of clusters. Some of the earliest work which allows overlapping clusters is based on defining a cluster as a locally optimal set (different locally optimal sets may overlap), [1], [2]. We refer to [4] for a survey on clustering.

The trend is to classify structure using global aggregate parameters (e.g. power laws) which emerge in the large scale limit. We tackle the opposite end of the spectrum, and ask:

*At what (small) scale does a network identify itself?*

A natural question would be "From what?" The natural answer is "From another network," but the intrinsic scale of one network should not depend on properties of another network. We propose a methodology which, given a network N with n vertices and m edges, extracts the intrinsic scale. Results from several networks reveal a surprising conclusion:

*The intrinsic scale of real networks is 7-20 vertices.*

Real networks betray their identity on a scale as small as a handful of notes. This means, these networks have non-trivial structure at small-scales, where aggregate parameters such as power-law exponents and clustering coefficients may not be stable.

A. Intrinsic Scale Via Distinguishability of Subgraphs Induced by Random Walks.

We argue that a network N has structure at scale κ if typical size-κ subgraphs from N are distinguishable from size-κ subgraphs in a randomized copy of N. This distinguishability implies "something" in N at scale κ must have been disturbed. Let N_δ be a perturbed copy of N with the same degree distribution, where δ quantifies the extent of the perturbation. In particular, $N_0 = N$ and N_∞ is a random graph with the same degrees as N. To construct perturbed graphs N_δ for $\delta = 1, 2, \ldots$, we use δ random edge-swaps to rewire the network. In an edge swap, edges (u, v) and (x, y) with distinct vertices u, v, x, y are rewired as follows:

$$(u, v) \quad (x, y) \quad \rightarrow \quad (u, x) \quad (v, y) \ ,$$
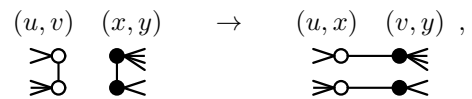
Observe that an edge swap preserves every vertex-degree. We illustrate a sequence of edge swaps on a toy graph below.
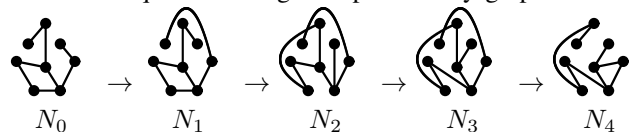
Random edge swaps slowly dismantle the "structure", yielding, in the limit, a random graph with the same degrees as N_0.

Let W be a random process that visits vertices. For concreteness, from now on W is a random walk which traverses a random incident edge at each step. After W visits κ different vertices, construct the subgraph induced by those κ vertices:

$$W : (N_\delta, \kappa) \mapsto G(\kappa, \delta), \qquad (1)$$

where $G(\kappa, \delta)$ is a random graph that depends on the network N_δ, the start vertex and edges traversed. The process W induces a distribution $p_{\kappa,\delta}$ over graphs with κ vertices. If the distributions $p_{\kappa,0}$ and $p_{\kappa,\delta}$ are distinguishable, existing structure in $N = N_0$ at scale κ was lost during the δ steps of randomization that produced N_δ. The Bayes *optimal* classifier for the distributions $p_{\kappa,0}$ and $p_{\kappa,\delta}$ has classification accuracy

$$\Delta(\kappa, \delta) = \tfrac{1}{2} \sum_G \max\{p_{\kappa,0}(G), p_{\kappa,\delta}(G)\}. \qquad (2)$$

We focus on $\delta \to \infty$, in which case, N_δ is a random graph with the same vertex-degrees as N. If $\Delta(\kappa, \infty) \gg \tfrac{1}{2}$, one can distinguish κ-sized subgraphs of N form those in N_∞ with high accuracy, which means there is significant structure at the scale κ in N. Hence, we define the intrinsic scale $\kappa^*(\tau)$:

Definition 1.1 (Intrinsic Scale): For $\tau > \tfrac{1}{2}$, let $\kappa^*(\tau, \delta)$ be the minimum scale κ at which one can distinguish κ-sized subgraphs of N from those in N_δ with accuracy at least τ,

$$\kappa^*(\tau, \delta) = \min\{\kappa \mid \Delta(\kappa, \delta) \geq \tau\}. \qquad (3)$$
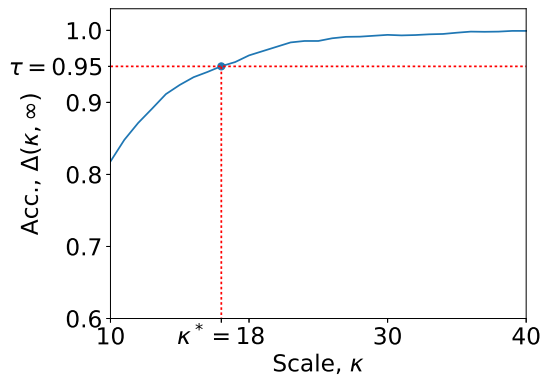
The intrinsic scale is $\kappa^*(\tau) = \lim_{\delta \to \infty} \kappa^*(\tau, \delta)$.
Implicit in the definition of intrinsic scale is the process W which produces κ-sized subgraphs. The details of W can affect specific values of κ^*, and it is natural to focus on subgraphs which are "locally" constructed as with a random walk.

Example (Intrinsic scale of trees): We show a 5-node tree N_0 in the figure below (leftmost). Edge swapping will randomly produce one of the 8 graphs shown (note, we allow parallel edges).
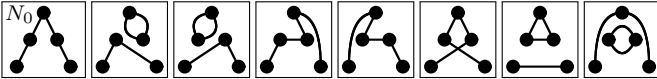
Connectivity distinguishes the original from the perturbed graph with 50% accuracy. For larger tree sizes, the accuracy improves.

With 95% accuracy, a random tree of size 18 can be distinguished from a random graph with the same degrees. The intrinsic scale at 95% accuracy is $\kappa^* \leq 18$ (upper bound

$\kappa^*(\tau)$	$\tau = 0.7$			$\tau = 0.9$			$\tau = 0.95$		
	30	50	∞	30	50	∞	30	50	∞
Road	8	7	**4**	24	16	**4**	34	22	**7**
Facebook	4	4	**4**	13	10	**8**	18	15	**10**
Human	28	21	**4**	*	*	**6**	*	*	**10**
Amazon	5	4	**4**	24	14	**9**	50	24	**12**
Al-Qaeda	4	4	**4**	14	10	**9**	21	15	**12**
Cite	8	6	**4**	32	20	**10**	64	33	**12**
DBLP	4	4	**4**	27	15	**9**	52	26	**13**
Web	7	5	**4**	36	21	**7**	64	37	**14**
Gowalla	16	11	**7**	*	*	**14**	*	*	**17**
Mouse	28	22	**4**	*	62	**15**	*	*	**20**
Yeast	28	18	**5**	*	*	**15**	*	*	**20**
Wiki	48	32	**9**	*	*	**16**	*	*	**20**

TABLE I: Intrinsic scale for $\delta \in \{30\%, 50\%, \infty\}$. We use $*$ to mean $\kappa^*(\tau, \delta) > 64$ (the maximum size in our experiments), which means that while there might be structure, it was robust to the perturbation (couldn't be systematically discriminated).

because we are not using the Bayes optimal classifier, just one based on connectivity). The example hits an important point. As the subgraph-size κ increases, computing the accuracy in [2] is exponential. To make the computation feasible, we summarized a subgraph using a statistic, connectivity, and obtained the classification accuracy using just that feature. This only gives a lower bound on the optimal accuracy. The same statistic may not work for every type of network. For example, with a large clique, random edge swaps would still maintain connectivity, and some other discriminative statistic would be needed to avoid the exponential complexity in [2].

Our notion of structure at scale κ corresponds to a game where I show you a random κ-sized subgraph and ask if you are surprised. You will be surprised if you see some "unexpected" structure. A 20-node clique might surprise *you* because you have an internal null distribution for random graphs, from which a 20-node clique is unlikely – has "unexpected" structure. We define this null distribution concretely as $p_{\kappa,\infty}$, which is natural as it is non informative over graphs with the same degrees. Our methodology, however, works with any other way to construct the null distribution while preserving desired properties of the graph (see for example [15]).

We summarize our main findings in Table I (the $\delta = \infty$ column), which gives upper bounds on the intrinsic scale of some real networks. Even at 95% accuracy, the intrinsic scale of real networks is no more than 20, for small and large networks alike. Traditionally structured networks, like roads, have smaller intrinsic scale (no surprise), while loosely structured networks like Wikipedia have larger intrinsic scale. Interestingly, the biological protein networks have comparatively large intrinsic scale, which indicates they have less structure than one might expect, perhaps due to the need for degeneracy, redundancy and robustness, [22]. At 70% accuracy, almost all networks have structure at very small scales that is fragile and easily disrupted with just 30% of edge-swaps.

Network	Type	# Nodes, n	# Edges, m
Road [13]	Infrastructure	1,088,092	1,541,898
Web [13]	Information	875,713	5,105,039
Amazon [11]	e-Commerce	334,863	925,872
DBLP [26]	Citation	317,080	1,049,866
Gowalla [3]	Social	196,591	950,327
Citation [7], [12]	Citation	34,546	421,578
Human [20]	PPI	8,077	26,085
Yeast [20]	PPI	5,718	48,253
Wiki [23], [24]	Information	4,604	119,882
Facebook [14]	Social	4,039	88,234
Mouse [20]	PPI	2,929	4,188
Al-Qaeda [10]	Social	271	756

TABLE II: Data networks used in our study.

Our Contributions. We propose a novel way to quantify structure in networks using an intrinsic scale at which identifying structures emerge and are distinguishable against a null distribution derived from random perturbations of the network. Real networks have structure at surprisingly small intrinsic scales. That structure, for small perturbation, can be fragile (easily dismantled) or robust.

II. DATA AND METHODS

We tested a variety of networks (see Table II for some information on the networks). All graph algorithms were implemented in Python using NetworkX [8]. The main challenge is to efficiently estimate the Bayes optimal accuracy $\Delta(\kappa, \delta)$ in [2], without computing the full distributions $p_{\kappa,0}$ and $p_{\kappa,\delta}$. Given $\Delta(\kappa, \delta)$, we compute the intrinsic scale using [3]. Our approach to computing $\Delta(\kappa, \delta)$ is to sample subgraphs and formulate the task as a standard machine learning problem. The workflow is as follows.

Algorithm 1: Estimate $\Delta(\kappa, \delta)$

1: Given N_0, construct N_δ using δ edge-swaps.

2: Sample κ-sized subgraphs from N_0 and N_δ to get a training set $\{G_{\kappa,0}\}^{\text{train}}$ and $\{G_{\kappa,\delta}\}^{\text{train}}$.

3: Use the training set to learn a classifier
$$g_{\kappa,\delta} : G_\kappa \mapsto \pm 1.$$
(+1 for N_0, −1 for N_δ).

4: Test the learned classifier $g_{\kappa,\delta}$ on independent test subgraphs $\{G_{\kappa,0}\}^{\text{test}}$ and $\{G_{\kappa,\delta}\}^{\text{test}}$.

5: Return $\hat{\Delta}(\kappa, \delta)$, the test accuracy of $g_{\kappa,\delta}$.

In step 1, edge-swaps preserve vertex degrees. For $\delta = \infty$, N_δ is a random graph with the same degrees as N_0. In steps 2 and 4, the training and test graphs are sampled using the random walker W. A larger training set gives a better learned classifier $g_{\kappa,\delta}$; a larger test set gives a better accuracy-estimate for $g_{\kappa,\delta}$. We used 10,000 samples from each graph, half for training and the rest for test. The Bayes optimal accuracy for the classification problem is $\Delta(\kappa, \delta) \geq \hat{\Delta}(\kappa, \delta)$. The best estimate of $\Delta(\kappa, \delta)$ comes from best learned classifier $g_{\kappa,\delta}$, hence the learning algorithm is important.

The hard task is in Step 3, which poses a graph classification problem. Any classifier trained in Step 3 gives an estimate $\hat{\Delta}(\kappa, \delta) \leq \Delta(\kappa, \delta)$. In [9], [25], a variety of approaches to graph classification are tested ranging from logistic regression and random forests using classical graph features (average degree, clustering coefficient, assortativity, etc.), to graph kernels, to deep convolutional networks (CNN) using lossless image representations of graphs proposed in [9], [25].[1] The best performing method is the CNN using the graph-image feature from [9], [25], and a close second is logistic regression on classical features. Choosing features is not easy, and can depend on the graph domain, hence we use the image representation in [9], [25] which is general and lossless. The CNN extracts appropriate features from this powerful graph image and learns a classifier. Using these graph images, Figure 1 illustrates how structure is perturbed with increasing edge-swaps for subgraphs from Facebook (a tightly structured network) and Wikipedia (a loosely structured network).

We make some qualitative observations from the pictures in Figure 1. Networks have signatures at different scales. As one perturbs a network, signature changes are visually discernible. Thus, a powerful CNN classifier using these graph image-signatures should come close to optimal classification accuracy. Further, different networks have different levels of structure at different scales. For example the Facebook signature at the 64-node scale is significantly disrupted by 10% edge-swaps, while the Wikipedia signature is not as disrupted. The level to which the signature at scale κ gets disrupted by δ edge-swaps is captured by $\Delta(\kappa, \delta)$, so we expect

$$\Delta_{\text{Facebook}}(64, 10\%) \gg \Delta_{\text{Wikipedia}}(64, 10\%).$$

At a small enough scale, the signature does not significantly change (e.g. the 16-node signatures in Figure 1). This suggests there is a critical scale κ^* at which the signature change becomes discernible with high accuracy.

Our experimental design is quite simple (see also Figure 2). For each network and for each pair of values (κ, δ), where

$$\kappa \in \{4, 5, 6, \ldots, 64\}$$
$$\delta \in \{10\%, 20\%, 30\%, 40\%, 50\%, \infty\},$$

we estimate $\Delta(\kappa, \delta)$ using $\hat{\Delta}(\kappa, \delta)$ from Algorithm 1. Note that δ is a percentage of the number of edges in the network, allowing us to compare networks of different sizes. We repeat each experiment for each network 10 times to reduce the variance due randomness in the construction of N_δ and the sampling of subgraphs to create training and test sets. In all cases, the learning algorithm is the CNN using the graph-image features, as already described earlier. For comparison, we also show some results for classifying based on topological graph features such as clustering coefficient and assortativity.

---

[1]In a nutshell, graph images are formed from the $\kappa \times \kappa$ adjacency matrix of a κ-node subgraph (1's are black pixels and 0s are white pixels). To structure the image into a signature which is invariant to isomorphism, one must order the vertices canonically, and the ordering which works best is based on a modified BFS with preference to high-degree nodes, see [9], [25] for details.
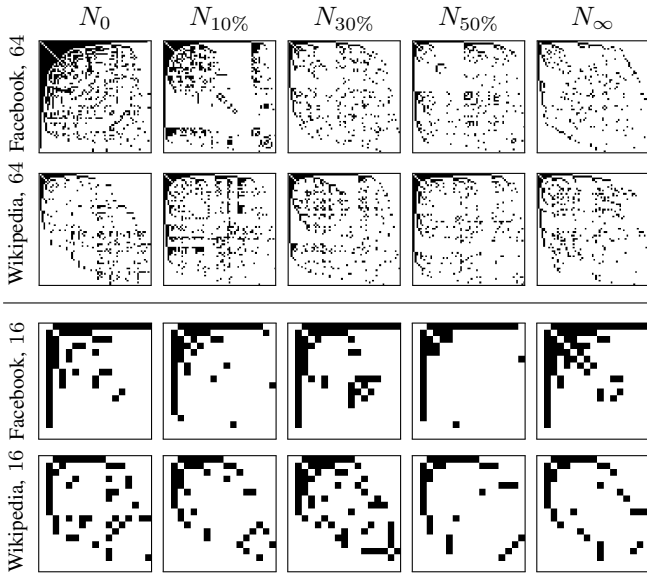
Fig. 1: **Network Signatures.** Leftmost are the "network signatures" (as pictures) of Facebook and Wikipedia at the 64-node and 16-node scales. Moving from left to right we show how that signature evolves as the network is perturbed with increasing number of edge-swaps δ, from 10% to ∞. For details on how these signatures (network pictures) are created we refer to [9], [25]. Rightmost are the network signatures of random graphs with the same vertex-degrees. At the 64-node scale, there is a clear change in signature from N_0 to N_∞, indicating that the "coordinated" structure has been disrupted. For Facebook, more so than Wikipedia, there is a significant change to the signature even for just 10% edge-swaps, which suggests Facebook is more "fragile" at this scale. At the 16-node scale, the signatures don't significantly change with increasing edge-swaps, suggesting that the structure at this scale is not coordinated enough to be disrupted by random edge-swaps.

III. EXPERIMENTS AND RESULTS

We show results on robustness and intrinsic scale.

A. Robustness

We first show results for small perturbations. Figure 3 shows the accuracy for just 10% edge-swaps. Recall that a high accuracy, $\Delta(\kappa) \gg \frac{1}{2}$, means the perturbed network at scale κ is highly distinguishable from the original network. This means structure in the network has been disrupted. Focusing on scale $\kappa = 24$ in Figure 3, we see that already at such a small scale, *for some networks*, there is significant distinguishability between the original network N and a 10%-perturbed copy of N. Indeed, the networks appear to cluster into three groups which we categorize loosely as fragile ($\Delta > 0.8$), semi-robust
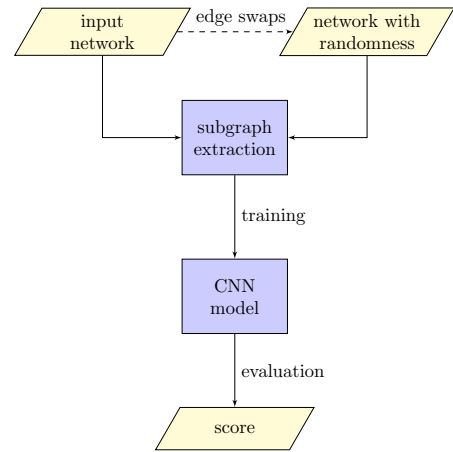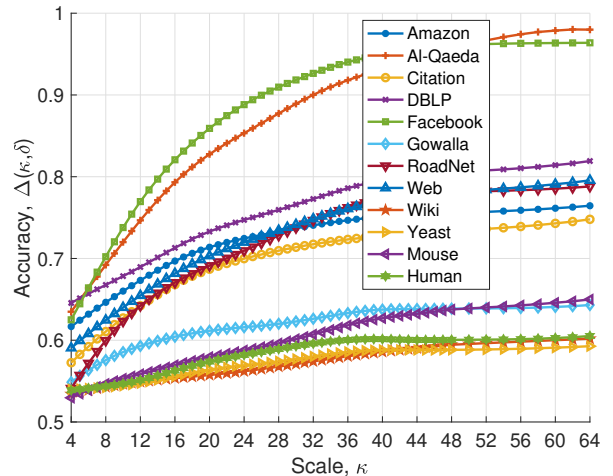

Fig. 2: Outline of experiments

Fig. 3: The accuracy $\Delta(\kappa, 10\%)$ for all networks with $\delta = 10\%$ edge-swaps. We cluster the networks by the accuracy of distinguishing N_0 from N_δ at scale $\kappa = 24$:

($\Delta \sim 0.7$) and robust ($\Delta < 0.6$).

fragile	($\Delta(24) > 0.8$)	Al-Qaeda; Facebook.
semi-robust	($\Delta(24) \sim 0.7$)	DBLP; Web; Road; Amazon; Citation.
robust	($\Delta(24) < 0.6$)	Human; Yeast; Wiki; Mouse; Gowalla.

In the fragile networks, which are the social networks, a small perturbation destroys the local structure leading to high distinguishability. This may not be a surprise as people usually choose their friends carefully and even small perturbations will disrupt those finely tuned social circles – this is especially so in the Al-Qaeda network which achieves more than 90% distinguishability with just 10% edge-swaps. In robust networks, the distinguishability with just a 10% perturbation is only marginally above random. This does not mean there is no structure at the 24-node scale. It just means the structure has
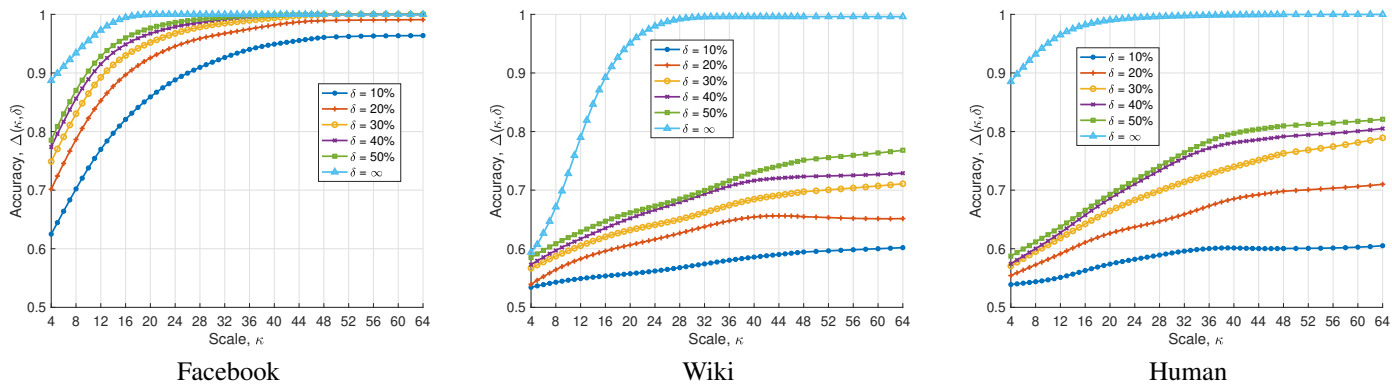
Fig. 4: Evolution of a network's structure as it is perturbed from $\delta = 10\%$ to ∞. A fragile network (Facebook) with intricate structure quickly becomes distinguishable with perturbation, even at small scale. A robust network (Wiki) resists the perturbations and remains somewhat indistinguishable from the original network until δ gets very large.

not yet been significantly disrupted by so small a perturbation. The biological networks fall into our classification of robust, which may indicate a level of redundancy/degeneracy that has been accumulated over the evolutionary process. The semi-robust networks are also interesting (DBLP, Web, Road, Amazon, Citation). These networks do have structure, but that structure is not so fragile as the social networks, indicating that the structure is not as fine tuned. Indeed, these networks have grown in an ad-hoc manner to represent the activity patterns of their actors, rather than being explicitly created by their actors (cf. social networks).

Figure 4 shows how structure gets dismantled as the perturbation increases from $\delta = 10\%$ to $\delta = \infty$ for three networks: Facebook (fragile social network); Wiki (robust information network); and, a biological network. Facebook quickly yields and after 50% edge-swaps the network has more-or-less reached a random graph with the same degrees. The Wiki network, on the other hand, resists, and even after 50% edge-swaps, the network is still not significantly discernible from the original unperturbed network. The mixing time for the edge-swapping random process is *much slower* on the robust Wiki network. The biological network resits small perturbations but slowly yields its structure with larger perturbations.

B. Intrinsic Scale

The view presented in Figure 4 highlights the evolution of a network as it is perturbed. Some networks vigorously resist even at large scales (hard to distinguish from the original network) and some fall apart even at smaller scales (easy to distinguish from the original network). We now go back to Figure 3 and focus on intrinsic scale. Figure 5 shows results analogous to Figure 3, but for increasing values of the perturbation δ. The typical behavior is a rapid rise in accuracy as scale increases, which corresponds to a rapid dismantling of the networks structure. This is followed by an elbow-turning point after which diminishing returns results in a flattening. The turning point (elbow) roughly corresponds to intrinsic scale, the scale at which all the observable structure has been

dismantled by the perturbation – going to larger scale does not improve accuracy significantly.

We now focus on $\delta = \infty$ to define the intrinsic scale. This choice of δ is to capture all the structure, whether fragile or robust – we must perturb hard enough to overcome the "robustness" of the network. For small perturbations, inability to distinguish the perturbed from the non-perturbed subgraphs may not indicate a lack of structure, but just that whatever structure exists may not yet have been dismantled. At $\delta = \infty$, all existing structure beyond the vertex degrees is gone. Indistinguishability now means there was no structure to start with. Distinguishability with high accuracy says that there was enough structure at the beginning. This structure may have been fragile or robust, but at $\delta = \infty$ we can't tell.

Visually looking at the elbows in Figure 5 for $\delta = \infty$ suggests that the networks roughly cluster into three categories.

Highly Structured	Road. (no surprise)
Semi-structured	Al-Qaeda, DBLP; Web; Road; Amazon; Citation; Gowalla.
Loosely structured	Wiki; Yeast; Mouse; Human.

Computing the elbow in the curves is not well defined and hard to generalize, so we opt for a simpler definition of intrinsic scale: the accuracy at $\delta = \infty$ must be above 95%. This accuracy threshold is quite strict and an intrinsic scale defined by the elbow will usually be smaller. Nevertheless, we opt for this simpler and more conservative definition. The intrinsic scales presented in Table I for different accuracy thresholds can all be obtained from Figure 5. The surprising conclusion is that for all these networks, spanning a variety of domains, the intrinsic scale is no more than 20 and as low as 7.

It is also interesting to note from Figure 5(c) that the accuracy approaches but doesn't quite reach 1. This asymptotic gap away from 1 indicates an amount of randomness in the original graph that cannot be distinguished from the random graph. This gap has about a 0.7 correlation with the intrinsic scale, and ranges from 0.17% for the Road network to about 0.58% for the Wiki network.
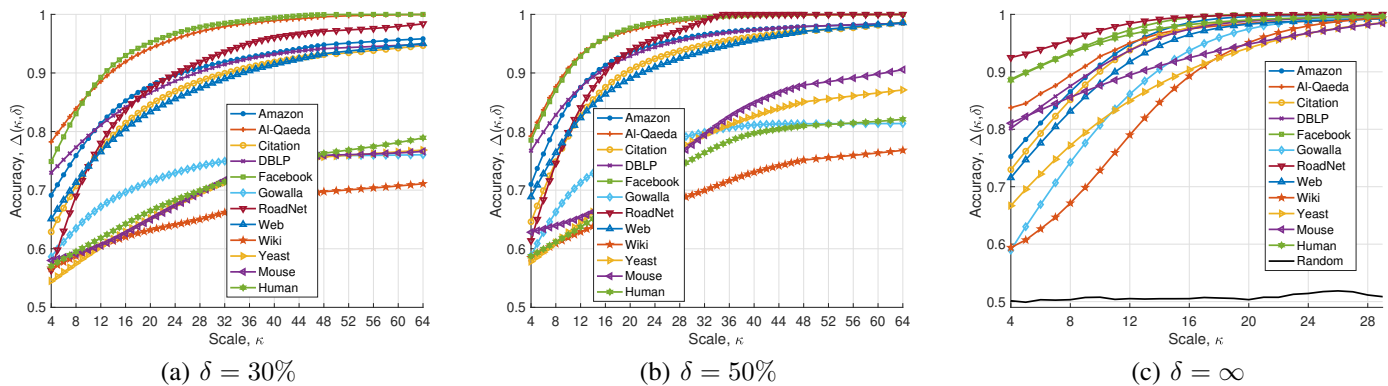
(a) $\delta = 30\%$ (b) $\delta = 50\%$ (c) $\delta = \infty$

Fig. 5: Intrinsic scale at different levels of perturbation δ. All networks display a "rapid" rise in accuracy (structure is rapidly lost) followed by an elbow followed by a flattening. The elbow could be considered the intrinsic scale at the particular value of perturbation δ. As a sanity check, we also show the accuracy for an Erdős-Renyi random graph at different scales for $\delta = \infty$. Such a graph has no "structure" at any scale, and it is no surprise that the accuracy hovers around $\frac{1}{2}$ for all scales.

C. Feature-Based Classification

Our algorithm to estimate $\Delta(\kappa, \delta)$ uses a learned classifier, and we have focused on the CNN with graph-images from [9], [25]. We briefly compare with more traditional feature-based methods. As a point of comparison, we take the Facebook network with $\delta = \infty$, and consider two classical features:

Clustering coefficient, C: Average fraction of closed triangles per vertex.
Measure of Assortativity, r: Average neighbor degree.

We show histograms of these features for 8-node subgraphs of the Facebook network and its perturbation in Figure 6.

The distributions $p_{8,0}$ and $p_{8,\infty}$ are clearly distinguishable. We can compute the Bayes optimal accuracy for each feature using [2] where the sum over graphs G is replaced by a sum of the feature's values. The results are in the table below.

Classifier	$\hat{\Delta}(8, \infty)$
Bayes optimal using C	0.905
Bayes optimal using r	0.820
Bayes optimal using C and r	0.932
CNN + graph-image	**0.934**

The CNN with the graph-image gives the best (highest) estimate $\hat{\Delta}$. Naturally, we can try other features and combinations of them, but one cannot exhaust all the possibilities for any given network, and further, a feature that works well for one type of network may not work well for another. And even still, there is no guarantee that the *optimal* estimate from using the features is better than the CNN plus graph-image. The graph-image feature is general, lossless and agnostic to the size and type of the network and when combined with the CNN gives top performance. Therefore CNN + graph-image was an easy choice for our classification problem.

D. Other Measures of Scale

Our intrinsic scale is not correlated with network-size (the correlations are negative: -0.6 with $|V|$, and -0.3 with

Network	Intrinsic scale	Cluster Size	Neigh. Size	Av. path length	Diam.
Road	7	5.95	2.83	308.91	753
Facebook	10	82.42	43.7	3.83	7
Human	10	12.26	6.46	4.25	7
Amazon	12	10.88	5.53	11.97	31
Al-Qaeda	12	8.47	5.58	3.5	4
Cite	12	48.45	24.4	4.36	10
DBLP	13	9.77	6.62	6.79	15
Web	14	19.08	11.7	6.34	16
Gowalla	17	17.65	9.67	4.62	11
Mouse	20	7.68	2.86	4.86	10
Yeast	20	18.56	16.9	3.28	6
Wiki	20	199.65	52.1	2.55	4
corr. coef.	1.000	0.3266	0.2257	-0.5047	-0.5043

TABLE III: Comparison of intrinsic scale with other measures of scale. The bottom row shows correlation of the other measures intrinsic scale.

$|E|$). We compare our measure of intrinsic scale with other reasonable measures of scale based on clusters and path-lengths.

Cluster size: Average of the cluster-sizes from the Speakeasy algorithm in [5].
1-neighborhood size: Also the average degree, $2|E|/|V|$.
Shortest path-length: Average over a large number of randomly sampled pairs of nodes.
Network diameter: A measure of global scale.

We compare our intrinsic scale with these other measures in Table III.[2] We also show the correlation coefficient of the other measures with intrinsic scale. None of the other measures are highly correlated with intrinsic scale. The closest is cluster size which can be much larger and dependent on the clustering algorithm. Intrinsic scale captures something non-trivial.

---

[2] Average path length and diameter are estimated from a sample of 10% of the vertex pairs.
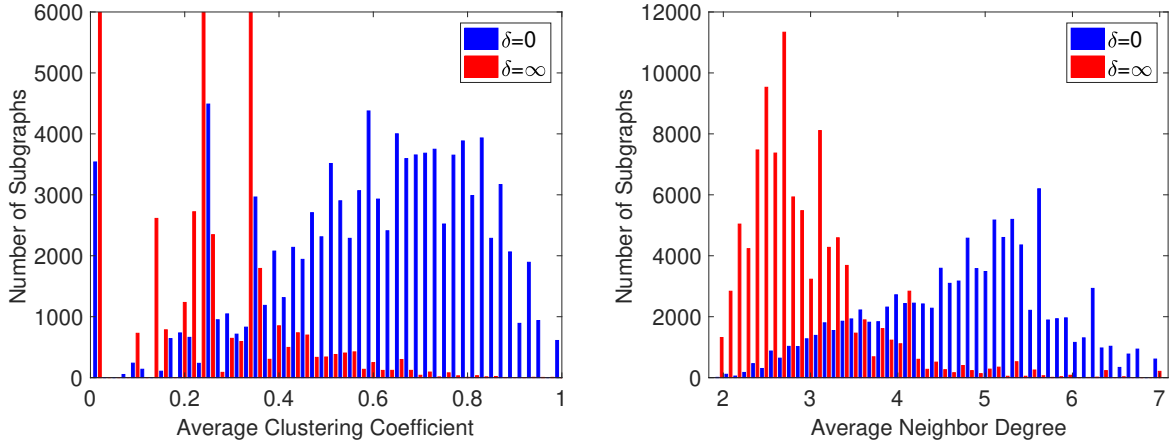
Fig. 6: Classical features: histograms of clustering coefficient (fraction of closed triangles) and assortativity (average neighbor degree) for 8-node subgraphs of the Facebook network with $\delta = 0$ and $\delta = \infty$. The perturbed and non-perturbed subgraphs induce distinguishable distributions over these features. However, quantitative analysis of these features produces a Bayes optimal classification error which is worse than the performance of the CNN plus image features.

IV. CONCLUSION AND FUTURE WORK

Our methodology for extracting the intrinsic scale of a network poses the task as a classification problem. This classification problem is to distinguish subgraphs on the network from subgraphs on a perturbed copy of the network. The accuracy $\Delta(\kappa, \delta)$ quantifies how much structure in the network at scale κ gets dismantled by a δ-perturbation. The learning curves for a fixed scale κ in Figure 7 show how the accuracy at that scale increases as one dismantles the structure in the network (by increasing δ). The rate at which structure gets dismantled for small perturbations is related to the robustness of the network, which we denote γ:

$$\text{robustness, } \gamma(\kappa) = -\ln(\Delta(\kappa, 10\%) - 0.5).$$

(logarithm(inverse of uplift in accuracy over random) for 10% perturbation). Robust networks hold on to their structure for small perturbations. For large perturbations, all the structure gets dismantled and the Bayes optimal accuracy quantifies the amount of structure there was in the network to start with, irrespective of robustness. We defined the intrinsic scale κ^* as the scale at which there is enough structure to achieve a classification accuracy exceeding 95%. A small intrinsic scale means the network is very structured.

We summarize our findings in Figure 8 which represents the networks in our study on a two-dimensional landscape of robustness and intrinsic scale. The social networks are especially fragile, and the biological networks are especially robust. One can approximately quantify the resilience of a network's functioning to vertex and edge removals using the degree-based parameter (see [6]):

$$\beta = \frac{\text{average squared degree}}{\text{average degree}}.$$

There is a moderate correlation of 37% between this measure of resilience β and our measure of robustness γ. A correlation
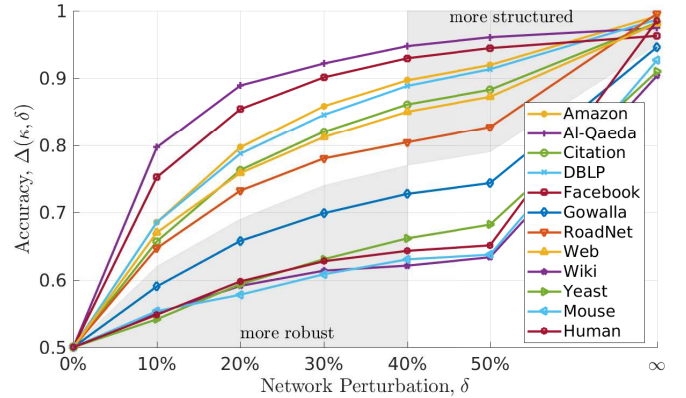

Fig. 7: Learning curves for $\kappa = 16$. Small perturbations reveal a network's robustness. Large perturbations, in particular $\delta = \infty$, reveal structure.

of 37% indicates some relationship between a network's ability to maintain its function under perturbation and the statistical recognizability of a networks topology against a null distribution obtained from a small (10%) perturbation. The relationship between statistical distinguishibility and resilience may warrant further study (akin to the relationship between statistical information and algorithmic compressability of sequences).

For the networks we examined, there is about a 61% correlation between structure and robustness. More structured networks with smaller intrinsic scale tend to be less robust. Our study provides a methodology for further investigation of this structure-robustness trade-off in networks. The trade-off is by no means universal: a notable exception is the Human PPI network which is very robust and yet very structured.

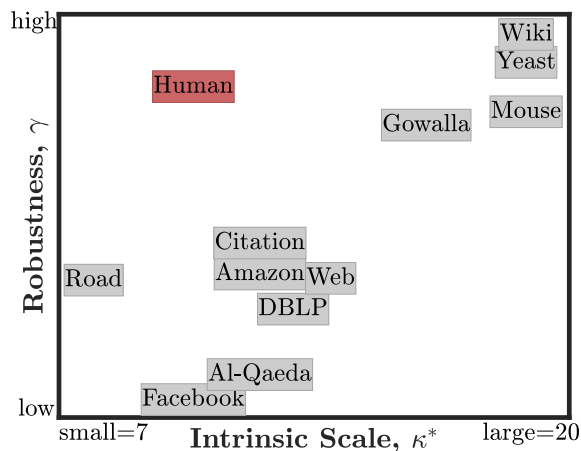Interesting future directions are: (i) Using statistical dis-

Fig. 8: Scatterplot of networks w.r.t. intrinsic scale and robustness. There appears a trend that more robust networks have higher intrinsic scale.

tinguishability, one can construct a taxonomy of real networks and random models with respect to the structure-robustness trade-off. One might then identify which models are appropriate for different real networks. (ii) How do we construct networks which break the structure-robustness trade-off, especially having very small intrinsic scale but very high robustness (e.g. Human PPI network). Such networks could have important applications. (iii) One can use knowledge about the intrinsic scale of a network to inform other network analysis algorithms such as clustering. For example, clusters should be defined with respect to information available within the intrinsic scale of the nodes participating in the cluster. The intrinsic scale can also guide the choice of hyperparameters in clustering algorithms which set bounds for cluster sizes, etc. Since the intrinsic scales of real networks are small, algorithmic analysis of such networks, when confined to scales on the order of the intrinsic scale, should be more efficient.

REFERENCES

[1] J. Baumes, M. Goldberg, and M. Magdon-Ismail. Efficient identification of overlapping communities. In *Proc. ISI*, 2005.
[2] Jeffrey Baumes, M. Goldberg, Mukkai Krishnamoorthy, M. Magdon-Ismail, and N. Preston. Finding communities by clustering a graph into overlapping subgraphs. In *Proc. IADIS*, 2005.
[3] Eunjoon Cho, Seth A Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. *KDD*, 2011.
[4] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75 – 174, 2010.
[5] Chris Gaiteri, Mingming Chen, Boleslaw Szymanski, Konstantin Kuzmin, Jierui Xie, Changkyu Lee, Timothy Blanche, Elias Chaibub Neto, Su-Chun Huang, Thomas Grabowski, et al. Identifying robust communities and multi-community nodes by combining top-down and bottom-up approaches to clustering. *Scientific Reports*, 5:16361, 2015.
[6] Jianxi Gao, Baruch Barzel, and Albert-Laszlo Barabasi. Universal resilience patterns in complex networks. *Nature*, 530(7590):307–312, 2016.
[7] Johannes Gehrke, Paul Ginsparg, and Jon Kleinberg. Overview of the 2003 KDD cup. *SIGKDD Newsl.*, 2003.
[8] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. *Proceedings of the 7th Python in Science Conference (SciPy2008)*, 2008.

[9] Kshiteesh Hegde, Malik Magdon-Ismail, Ram Ramanathan, and Bishal Thapa. Network signatures from image representation of adjacency matrices: Deep/transfer learning for subgraph classification. *arXiv:1804.06275*, 2018.
[10] JJATT. John jay & artis transnational terrorism database, 2009.
[11] Jure Leskovec, Lada A Adamic, and Bernardo A Huberman. The dynamics of viral marketing. *TWEB*, 2007.
[12] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: Densification laws, shrinking diameters and possible explanations. *KDD*, 2005.
[13] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Math.*, 2009.
[14] Jure Leskovec and Julian J. Mcauley. Learning to discover social circles in ego networks. *NIPS*, 2012.
[15] Sach Mukherjee and Terence P. Speed. Network inference using informative priors. *Proceedings of the National Academy of Sciences*, 105(38):14313–14318, 2008.
[16] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.
[17] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, 2004.
[18] E. Olbrich, T. Kahle, N. Bertschinger, N. Ay, and J. Jost. Quantifying structure in networks. *The European Physical Journal B*, 2010.
[19] Ray Reagans and Bill McEvily. Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48(2):240–267, 2003.
[20] Jüri Reimand, Laur Tooming, Hedi Peterson, Priit Adler, and Jaak Vilo. Graphweb: mining heterogeneous biological networks for gene modules with functional significance. *Nucleic acids research*, 36(suppl_2):W452–W459, 2008.
[21] Stephen B. Seidman. Network structure and minimum degree. *Social Networks*, 1983.
[22] Giulio Tononi, Olaf Sporns, and Gerald M. Edelman. Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences*, 96(6):3257–3262, 1999.
[23] Robert West and Jure Leskovec. Human wayfinding in information networks. *WWW*, 2012.
[24] Robert West, Joelle Pineau, and Doina Precup. Wikispeedia: An online game for inferring semantic distances between concepts. *IJCAI*, 2009.
[25] Ke Wu, Philip Watters, and Malik Magdon-Ismail. Network classification using adjacency matrix embeddings and deep learning. *ASONAM*, 2016.
[26] Jaewon Yang and Jure Leskovec. Defining and evaluating network communities based on ground-truth. *ICDM*, 2012.