Dense Graphs and Sparse Matrices

Milan Randić* and Luz M. DeAlba

Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311

Received July 2, 1997⊗

We consider rigorous definitions for dense graphs and sparse matrices, thus quantifying these concepts that have been hitherto used in a qualitative manner. We assign to every graph the compactness index ρ, which is a measure of density of graphs and binary matrices. This index takes values $\rho > 1$ for dense graphs and $\rho < 1$ for sparse graphs (matrices). The index is derived from the quotient of the relative graph density and the percentage of zero entries in the adjacency matrix. The numerical values for the compactness index are reported for miscellaneous graphs, including several families of structurally related graphs.

INTRODUCTION

In graph theory occasionally one speaks of dense graphs, just as one speaks of sparse matrices in matrix theory, without specifying precisely what dense graphs or sparse matrices are. For example, in discussing *arboricity* of graphs (defined as the minimal number of line-disjoint spanning forests of a graph) F. Harary[1] refers to K_5 as a dense graph. Similarly D. B. West[2] speaks of dense graphs in a discussion of random graphs with a linear number of edges and constant probability (i.e., independent of n, the size of the graph). In both cases, clearly, the attribute "dense" is used in a qualitative manner since neither case gives a definition for dense graphs. Examination of a dozen books on graph theory also failed to locate a rigorous definition of dense graphs. The situation is the same with respect to sparse matrices. According to Harary (ref 1 footnote on p 205) "In the literature, a sparse matrix has been defined as one with many zeros". How many?

Can dense graphs and sparse matrices be rigorously defined? We feel that they can, even though there will be graphs that approach the border dense/sparse for which the classification may appear not so clearly marked. That is, there may be graphs of similar appearance one of which will be classified as dense and the other as sparse. But that is also the case with quantities that are continuous by nature. For example, where in the spectrum does green color begin and where does it end? This is not so clear unless we specify which wavelength belongs to which color. Graphs are discrete objects and from that point of view the task of classification of graphs as dense or sparse appears less challenging, but as we will see the solution is not straightforward if one want to avoid ambiguities associated with arbitrariness in classifying graphs at the border line.

Such problems do not arise when graphs are classified according to well defined structural elements defined by the presence or absence of the critical structural elements. In Table 1 we list several well-known classifications of graphs. To this list we added the recently introduced notion of saturation.[3] The property of saturation is associated with the detour matrix of graphs.[1,3] The detour matrix records the lengths of the longest paths for each pair of vertices in a graph. In contrast to other graph matrices nonisomorphic graphs can have identical detour matrix for suitable numbering of the vertices. Saturated graphs are graphs which have the detour matrix the same as the complete graph on n vertices. For saturated graphs all of the entries in their detour matrix have attained the maximal value possible, which for a graph having n vertices is $n - 1$. Hence each pair of vertices in a saturated graph is connected by a Hamiltonian path, and addition of edges cannot increase the maximal distance between any two vertices.

Table 1. Classification of Graphs

planar/nonplanar	bipartite	isospectral
cyclic/acyclic	polyhedral	endospectral
transitive	n-connected	cages
vertex transitive	cubic	maximally planar
edge transitive	complete	hypercubes
eulerian	complete bipartite	saturated
Hamiltonian	zero symmetric	maximally saturated

The notion of maximally saturated graphs offers an opportunity to classify the graphs which have the same number of vertices and a different number of edges into two classes: (1) graphs which have the detour matrix of K_n, and (2) graphs which have not yet attained the maximal entries in their detour matrix. However, the above classification does not parallel the usual notions of dense graphs and sparse matrices. In Figure 1 we illustrate a family of structurally related graphs of increasing size, the detour matrix of which is maximal.[3] By adding an edge to any of these graphs their detour matrix does not change. By subtracting any of the edges in any of the graphs of Figure 1 the detour matrix will change because at least for one pair of vertices the longest path will shorten. Hence, these graphs represent critical graphs with respect to the saturation. Yet the corresponding adjacency matrices, as the size of the graphs increases, become apparently more and more sparse, or less and less dense. Therefore the detour matrix and the saturated graphs, while well defined, do not offer a route for differentiation between dense graphs and sparse graphs (matrices).

ON DENSITY OF GRAPHS

The density in graphs has been defined as the ratio E/V [ref 2, Glossary of Terms, p 457]. This ratio though well-defined does not offer a useful basis for the definition of dense graphs. For example, all cubic graphs have the same ratio E/V yet small cubic graphs appear dense, while the adjacency matrix of large cubic graphs is apparently sparse. Clearly, the definition of *dense* and *sparse* ought to reflect the relative number of zero and nonzero matrix elements in the adjacency matrix of a graph.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1997.

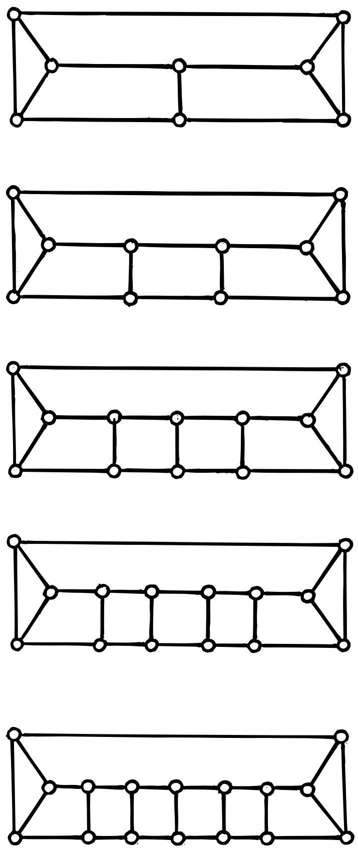

Figure 1.

Table 2. The Density (E/V), the Relative Density (E/E^*), the Quotient of the Number of Zeros and the Number of Ones in the Adjacency Matrix ($(n^2 - 2E)/n^2$), and ρ, the Quotient $(E/E^*)/((n^2 - 2E)/n^2)$ for the Saturated Graphs of Figure 1

n	E/V	E/E^*	Z/Z^*	ρ
6	3/2	3/5	18/36	1.200 00
8	3/2	3/7	40/64	0.685 71
10	3/2	3/9	70/100	0.476 19
12	3/2	3/11	108/144	0.363 64
	...	...	...	...
limit	3/2	0	1	0

An alternative measure of the density of graphs is the quotient E/E^*, where E is the number of edges and E^* is the number of edges in the complete graph having the same number of vertices. The quotient E/E^* has been considered in the literature.[4] We will refer to the quotient E/E^* as the relative density of a graph, since it can be viewed as the ratio of density of a graph to the density of the complete graph having the same number of vertices, i.e., $(E/V)/(E^*/V)$.

In Figure 1 we show a family of trivalent (cubic) "prismatic" graphs, and in Table 2 we list their alternative density measures. As we see the relative density of cubic graphs is not constant but decreases as the graphs become larger. Hence, the quotient E/E^* differentiates more dense from less dense graphs. However, the quotient E/E^* has an apparent conceptual disadvantage: Once the critical value of E/E^* has been selected the question remains should graphs with the critical value for E/E^* be viewed as dense or sparse. For example, if we select the value $E/E^* = 1/2$ as the critical value that discriminates dense from sparse, should graphs with the value $E/E^* = 1/2$ be classified as dense or sparse? Any such choice would be inherently equally arbitrary. This problem is typical of other classification schemes in which the objects to be classified can attain the critical value used for the classification.

Let us consider an alternative index of density of graphs. The quotient Z/Z^*, where Z is the number of zero entries in the adjacency matrix of a graph and Z^* is the number of entries in the adjacency matrix (or the number of zeros for null-graph!), gives the percentage of zero entries in the graph. For the graphs of Figure 1 we show in Table 2 the magnitudes of the quotient Z/Z^*. The quotient varies with the size of cubic graphs and increases as the size of the graphs increases. Here we could select the value $Z/Z^* = 1/2$ (i.e., 50%) as the critical measure of the dense/sparse character of a graph. Again we encounter the same difficulty in deciding whether the graphs with $Z/Z^* = 1/2$ should be considered dense or sparse.

The two outlined measures of the dense/sparse character of a graph E/E^* and Z/Z^* do not fully overlap, although they inversely parallel to each other. According to the E/E^* index, the first graph in Figure 1 is dense, while the Z/Z^* index places it at the border, i.e., it could be viewed as dense or sparse, depending on whether we decide $Z = 1/2$ will represent dense or sparse graphs, respectively.

DESIDERATA

What is desired is an index of dense/sparse characterization of graphs which will (1) parallel the intuitive notion of dense and sparse graphs while (2) at the same time define, unambiguously, the critical boundary that separates dense and sparse graphs so that no graph can be at the boundary. Is this possible? Can we define a critical value numerically so that every graph has either a greater or smaller value than the critical value? If this can be done, we would not have the problem mentioned previously in arbitrary selecting one or the other alternative that such considerations introduce.

The indices E/E^* and Z/Z^* satisfy the first requirement; they apparently follow the intuitive notion of dense and sparse graphs. However, both indices fail to offer a critical value for the index that separates dense from sparse graphs so that no graph assumes the critical value. In the next section we will outline an index that satisfies the second, obviously a more difficult, requirement. We will refer to such an index of dense/sparse characters as the graph compactness, to avoid confusion with terms "dense" and "density" that have already been used.

GRAPH COMPACTNESS

Consider the double quotient: $(E/E^*)/(Z/Z^*)$. Let us again focus attention on the cubic graphs of Figure 1 and examine the behavior of the double quotient for the case of large graphs. As the graphs increase in size, the quotient Z/Z^* approaches 1, and therefore E/E^* measures the variation in the graph compactness. On the other hand, if we consider very dense graphs, graphs that approach the complete graph K_n, then the first quotient E/E^* approaches 1, and the quotient Z/Z^* measures of the variation in the graph compactness. Hence, both measures of the dense/sparse character appear suitable for extreme situations. This suggests the double quotient $(E/E^*)/(Z/Z^*)$ as a potentially suitable index of graph compactness. The two factors E/E^* and Z^*/Z (the reciprocal of Z/Z^*) parallel each other, so their product, which is the double ratio, to be designated as ρ, will preserve the properties of both indices.

We will show now that the new index satisfies both conditions that we put forward as desirable for a measure of dense/sparse character of graphs. For a graph having n vertices the quotient E/E^* can be written as $E/n(n-1)$, there being $n(n-1)$ nonzero entries above the diagonal of the $n \times n$ adjacency matrix of the complete graph K_n. The quotient Z/Z^* for a graph having n vertices and E edges can be expressed as $Z = (n^2 - 2E)/n^2$. Therefore the ratio of the two quotients becomes

$$\rho = (n^2/2\mathrm{E} - 1)(1 - 1/n)$$

This ratio can take values greater and smaller than 1. The value $\rho = 1$ appears a natural critical value that separates dense and sparse graphs. Graphs with $\rho > 1$ (graphs which have few zeros) are classified as dense, and graphs with $\rho < 1$ (matrices which have many zeros) are classified as sparse.

It is more difficult to satisfy the second condition, but our index ρ has past the test. If ρ is to be 1 then

$$(n^2/2\mathrm{E} - 1) = n/(\mathrm{n} - 1)$$

Both n and E can be viewed as free variables. We can solve the above for E obtaining

$$E = n^2(\mathrm{n} - 1)/2(4n - 1)$$

E has to be an integer so the question is if the above equation has a solution. The right hand side, the fraction, can be an integer only if the denominator is a factor of the numerator, but this clearly is not the case since the numerator is fully factored and shows different factors. Hence there is no graph that can have $\rho = 1$. We can therefore propose the following definitions for dense graphs and sparse matrices, respectively.

Definition 1. Dense graphs (dense matrices) are graphs (matrices) with the compactness ratio ρ that is less than the critical value $\rho = 1$. Here ρ is the ratio of quotients: Z/Z^*, the ratio of the zeros to the total number of entries in the adjacency matrix, and the ratio E/E^*, of the relative density of the graph and K_n.

Definition 2. Sparse matrices (sparse graphs) are matrices (graphs) with the ratio ρ that is less than the critical value $\rho = 1$, where ρ is the ratio of two quotients: the number of the zeros to the total number of entries in the adjacency matrix and the ratio of the relative density of the graph and K_n.

In Table 2 we list the quotients $(n^2 - 2E)/n^2$ and $2E/n(n - 1)$ and the double quotient ρ for the saturated graphs of Figure 1. As we see, except for the first graph of Figure 1, the graph of the trigonal prism, all other saturated graphs of Figure 1 are classified as sparse. As their size increases so also the dense/sparse character decreases and in the limit approaches zero. In other words we can construct as sparse a graph of this type as we want if we increase the size of the graphs.

ILLUSTRATIONS

It is easy to show that the compactness index for complete graphs K_n equals n. This simple proportionality of the compactness index with the size index reflects the fact that it takes more bonds to be broken to dismember the complete graph—hence a justification for the label "compactness". In Table 3 we list the densities, the relative densities, the Z/Z^* index, and the value of ρ for complete bipartite graphs $K_{n,n}$.

Table 3. The Density (E/V), the Relative Density (E/E^*), the Quotient of the Number of Zeros and the Number of Ones in the Adjacency Matrix ($(n^2 - 2E)/n^2$), and ρ, the Quotient $(E/E^*)/((n^2 - 2E)/n^2)$ for the Complete Bipartite Graphs $K_{n,n}$

graph	E/V	E/E^*	Z/Z^*	ρ
$K_{1,1}$	1/2 = 0.5	1	1/2	2.000 000 0
$K_{2,2}$	4/1 = 1	4/6 = 2/3	1/2	1.333 333 33
$K_{3,3}$	9/6 = 1.5	9/15 = 3/5	1/2	1.200 000 00
$K_{4,4}$	16/8 = 2	16/28 = 4/7	1/2	1.142 857 14
$K_{5,5}$	25/10 = 2.5	25/45 = 5/9	1/2	1.111 111 11
$K_{6,6}$	36/12 = 3	36/66 = 6/11	1/2	1.090 909 09
$K_{7,7}$	49/14 = 3.5	49/91 = 7/13	1/2	1.076 923 08
$K_{8,8}$	64/16 = 4	64/120 = 8/15	1/2	1.066 666 67
...				...
limit $K_{n,n}$	$2n$	n/(2n − 1)	1/2	1.000 000 0

Table 4. The Density (E/V), the Relative Density (E/E^*), the Quotient of the Number of Zeros and the Number of Ones in the Adjacency Matrix (Z/Z^*), and ρ, the Quotient $(E/E^*)/(Z/Z^*)$ for Polyhedral Graphs[a]

graph	V	E/V	E/E^*	(Z/Z^*)	ρ
Sparse Graphs					
truncated icosahedron	60	1.5	90/1770	3420/3600	0.048 305
truncated cube	24	1.5	36/276	504/576	0.149 068
truncated octahedron	24	1.5	36/276	504/576	0.149 068
snub cube	24	2.5	60/276	456/576	0.172 101
dodecahedron	20	1.5	30/190	360/400	0.175 439
small rhombicuboctahedron	24	2	48/276	480/576	0.208 670
pentagonal prism	10	1.5	15/45	70/100	0.233 333
cuboctahedron	12	2	24/66	96/144	0.242 424
truncated tetrahedron	12	1.5	18/66	120/144	0.327 723
rhombic dodecahedron	14	1.7143	24/91	148/196	0.349 272
trapezohedron	12	1.6667	20/66	104/144	0.419 580
triakis octahedron	14	2.5714	36/91	124/196	0.625 310
tetrakis octahedron	14	2.5714	36/91	124/196	0.625 310
cube	8	1.5	12/28	40/64	0.685 714
pentagonal antiprism	10	2	20/45	60/100	0.740 740
icosahedron	12	2.5	30/66	84/144	0.779 221
Dense Graphs					
square antiprism	8	2	16/28	32/64	1.142 857
trigonal bipyramid	5	1.8	9/10	16/25	1.406 250
triakis tetrahedron	8	2.25	18/28	28/64	1.469 388
hexagonal bipyramid	8	2.25	18/28	28/64	1.469388
octahedron	6	2	12/15	12/36	2.400 000
tetrahedron	4	1.5	6/6	4/16	4.000 000

[a] The figures of the polyhedral graphs considered can be found in ref 5.

As we see from Table 2 the density E/V increased with n, and E/E^* slowly decreases while Z is constant. The ρ values for the complete bipartite graphs $K_{n,n}$ decreases from 2 (the value of $K_{1,1}$) to 1, the limiting value for an infinitely large complete bipartite graphs. Hence, all complete bipartite graphs $K_{n,n}$ are dense, though as their size increases their dense/sparse character decreases.

In Table 4 we listed the densities and the dense/sparse character for polyhedral graphs. Most polyhedral graphs represent sparse graphs. Only rather small polyhedra, like the Archimedean square antiprism, switch over the classification and qualify as a dense graph.

In Table 5 we show the compactness for miscellaneous graphs that occurred in the literature on graph theory. As we see in each case the classification dense/sparse agrees with the intuitive notion of sparse matrices and dense graphs. Thus the complete graph K_5 (and other complete graphs) is dense, in agreement with the statement of Harary. Among dense graphs we find bipartite complete graphs $K_{n,n}$. The

Table 5. The Density (E/V), the Relative Density (E/E^*), the Quotient of the Number of Zeros and the Number of Ones in the Adjacency Matrix (Z/Z^*), and ρ, the Quotient $(E/E^*)/(Z/Z^*)$ for Miscellaneous Graphs

Sparse Graphs					
Desargues−Levi graph	20	1.5	30/190	360/400	0.175 439
Blanusa graph	18	1.5	27/153	270/324	0.211 765
Heawood graph	14	1.5	21/91	154/196	0.293 706
Petersen graph	10	1.5	15/45	70/100	0.476 190
Dense Graphs					
Harary graph $H_{5,8}$	8	2.5	20/28	24/64	1.904 762
Kuratowski $K_{3,3}$	6	1.5	9/15	18/36	1.200 000
Kuratowski K_5	5	2	10/10	15/25	1.666 667

adjacency matrices of polyhedral cubic graphs, including hypercubes (with exception of the tetrahedron which is represented by K_4), are sparse.

Some properties of dense graphs have been mentioned in the literature without specifying the class of dense graphs. For example, if a graph is sufficiently dense it will have a Hamiltonian cycle. It remains to be seen if our definitions of sparse matrices and dense graphs can help to "translate" such qualitative statements involving density of graphs to rigorous quantitative statements. That is clearly outside the scope of the work presented here.

CONCLUDING REMARKS

Until now the attributes "dense" and "sparse" have been used for graphs and matrices in a qualitative manner. We proposed definitions for the two concepts that allows one to assign a quantitative measure to these quantities. Clearly, these are our definitions, the first definitions yet proposed for these concepts, and users will judge whether they are useful and how useful they are. Others may propose alternative definitions, but what is more likely is that some may argue that one does not need rigorous definitions for these concepts that have been widely used only in a qualitative manner. To that attitude we can reply with a quote from Lord Kelvin:[6] "When you can measure what are you speaking about, and express it in numbers, you know something about it. But when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind: It may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science".

REFERENCES AND NOTES

(1) Harary, F. *Graph Theory*; Addison-Wesley: Reading, MA, 1969.
(2) West, D. B. *Introduction to Graph Theory*; Prentice Hall: Upper Saddle River, NJ, 1966.
(3) Randić, M.; DeAlba, L. M.; Harris, F. E. Graphs with the same detour matrix. *Croat. Chem. Acta* In press.
(4) Roberts, F. S. *Applied Combinatorics*; Prentice-Hall Inc.: Englewood Cliffs, NJ, 1984.
(5) Williams, R. *The Geometrical Foundation of Natural Structure*; Dover Publ. Inc.: New York, 1979.
(6) Lord Kelvin, *Popular Lectures & Addresses 1891−1894*.

CI970241Z