

Buscemi Centrality: Source-Relative Centrality in Heterogeneous Affiliation Graphs

T. H. Underpoot¹

Zeech Flugelhorn¹

¹Centre for Unnecessary Formalisation, [redacted]

Abstract

This formulation enables the incorporation of heterogeneous interaction semantics, path-dependent evidential strength, and localised structural context into a unified source-relative score that is robust to sparsity, resistant to trivial shortest-path degeneracies, and broadly applicable across multimodal affiliation networks or whatever.

Distance-to-source metrics such as Erdős numbers provide an intuitive measure of proximity to distinguished individuals within affiliation graphs. However, these metrics collapse structural information to a single shortest-path value and do not account for heterogeneity in interaction types or variation in evidential strength.

We introduce *Buscemi centrality*, a source-relative centrality measure defined on heterogeneous graphs with typed edges and path-quality semantics. The measure combines (i) best-path accessibility to a distinguished source under interaction-dependent quality and cost, with (ii) contextual embeddedness within the source’s local neighbourhood. This formulation captures both proximity and interaction semantics while remaining distinct from distance-aggregation centrality measures.

We demonstrate that nodes with identical source-distance may exhibit substantially different Buscemi centrality, reflecting differences in interaction quality and structural embedding. We discuss computational considerations and relate the formulation to distance-based centrality frameworks, including harmonic centrality as studied by Boldi and Vigna.

1 Introduction

Graph-based proximity measures to distinguished individuals have long been used as informal indicators of structural position. The Erdős number is

a canonical example defined on coauthorship graphs, with analogous constructions in other domains, such as the well-known Bacon number in film coappearance networks.

Such measures are appealing due to their simplicity but exhibit two limitations. First, they depend solely on a single shortest path. Second, they assume homogeneous edge semantics, ignoring qualitative differences between interaction types.

In many real-world systems, relationships are heterogeneous, with interactions differing in strength, reliability, and semantic meaning. Such systems are naturally modelled as weighted or attributed graphs [4]. Direct interpersonal contact, collaboration, and co-membership represent different forms of association with different evidential strengths. Treating these as identical edges can obscure meaningful distinctions.

In this work, we introduce Buscemi centrality,¹ a source-relative centrality measure defined on heterogeneous affiliation graphs. The measure incorporates both path quality and local structural context, enabling distinctions between nodes that are indistinguishable under classical distance-based metrics.

2 Graph Model

Let $G = (V, E, \tau)$ be a graph where:

- V is a set of individuals,
- $E \subseteq V \times V$ is a set of edges,
- $\tau : E \rightarrow \mathcal{T}$ assigns each edge a type.

Each edge type $t \in \mathcal{T}$ represents a distinct interaction category (e.g. coauthorship, coappearance, documented contact), extending standard weighted graph formulations to heterogeneous interaction models [4].

We associate each type t with:

- a traversal cost $c(t) \geq 0$,
- a quality factor $q(t) \in (0, 1]$.

Costs reflect interaction distance, while quality factors reflect evidential strength.

¹We name this measure after Steve Buscemi. He has not been informed.

3 Path Quality and Accessibility

For a path $P = (v_0, \dots, v_k)$, define its quality:

$$Q(P) = \frac{\prod_{i=0}^{k-1} q(\tau(v_i, v_{i+1}))}{1 + \sum_{i=0}^{k-1} c(\tau(v_i, v_{i+1}))}. \quad (1)$$

For nodes $u, v \in V$, define accessibility:

$$A(u, v) = \max_{P: u \rightsquigarrow v} Q(P), \quad (2)$$

with $A(u, v) = 0$ if no path exists. We additionally define $A(v, v) = 1$ for all $v \in V$, corresponding to the empty path of zero cost and unit quality, isn't it.

We consider only simple paths in the maximisation above.² This excludes degenerate cycles, which cannot strictly increase $Q(P)$ under the constraints $q(t) \leq 1$ and $c(t) \geq 0$; in the limiting case $q(t) = 1$ and $c(t) = 0$, cycle traversal leaves quality unchanged.

The functional form of $Q(P)$ balances multiplicative accumulation of evidential strength with additive penalisation of traversal cost, ensuring that longer, weaker paths do not dominate while still allowing strong interaction chains to outweigh shorter but less informative ones.

4 Source Neighbourhood

Let $s \in V$ be a distinguished source node.

Define the weighted distance $d_w(u, s)$ using traversal costs $c(t)$.

For a fixed radius r , define the source neighbourhood:

$$N_r(s) = \{u \in V : d_w(u, s) \leq r\}. \quad (3)$$

Define neighbourhood weights:

$$\alpha(u; s) = \frac{A(u, s)}{\sum_{x \in N_r(s)} A(x, s)}. \quad (4)$$

These weights reflect how strongly each node belongs to the source's local structure.

²In principle one could traverse a cycle indefinitely. We have chosen not to.

5 Buscemi Centrality

We define:

$$BC(v; s) = \lambda A(v, s) + (1 - \lambda) \sum_{u \in N_r(s)} \alpha(u; s) A(v, u), \quad (5)$$

where $\lambda \in [0, 1]$.

The parameter λ interpolates between two limiting cases: $\lambda = 1$ yields $BC(v; s) = A(v, s)$, while $\lambda = 0$ yields a purely neighbourhood-based measure reflecting embeddedness within $N_r(s)$.

We observe that a measure named *Buscemi* centrality admits only one sensible choice of source node. We therefore fix $s = \text{Buscemi}$ for the remainder of this work and introduce the simplified notation:

$$BC(v) := BC(v; \text{Buscemi}), \quad (6)$$

$$N_r := N_r(\text{Buscemi}), \quad (7)$$

$$\alpha(u) := \alpha(u; \text{Buscemi}). \quad (8)$$

QED, n'est ce pas?

6 Relation to Prior Work

Distance-based centrality measures, including harmonic centrality, aggregate reciprocal shortest-path distances across a node set. Boldi and Vigna [2] provide an axiomatic treatment.

Buscemi centrality is not reducible to such measures. It depends on typed edge semantics, non-linear path-quality aggregation, and source-conditioned neighbourhood weighting, none of which can be expressed purely in terms of shortest-path distances.

Nodes with identical shortest-path distance to Buscemi may therefore exhibit arbitrarily different Buscemi centrality values.

6.1 Relation to Bacon Numbers and Erdős Numbers

Consider the degenerate case in which all edge types are assigned $q(t) = 1$ and $c(t) = 1$. Then for a path of d hops, path quality reduces to $Q(P) = 1/(1 + d)$. Accessibility becomes $A(v) = 1/(1 + d_{\min}(v))$ where $d_{\min}(v)$ is the minimum hop count to Buscemi, and the ranking induced by $A(v)$ is equivalent to the ranking induced by the Bacon number with source set to Steve Buscemi.

The Bacon number is itself a special case of the Erdős number construction applied to a film coappearance graph with a different distinguished source. Buscemi centrality therefore strictly generalises both the Bacon number and the Erdős number.

It follows that all of graph theory is, at some level, secretly about Steve Buscemi. Some may argue that this is a non-sequitur, but the authors maintain that this is axiomatic.³

7 Computation

Accessibility values $A(u, v)$ can be computed using modified shortest-path procedures over simple paths. While the multiplicative term admits logarithmic transformation, the denominator introduces global path dependence, preventing reduction to a standard shortest-path problem.

In practice, label-setting or label-correcting methods can track Pareto-optimal path candidates. For source-relative computation, $A(\cdot, \text{Buscemi})$ is obtained via a single-source traversal from Buscemi.

For large-scale graphs, approximate neighbourhood computation techniques such as HyperANF [1] may be used, with pruning strategies to control the number of candidate paths. Such approximations are particularly relevant for large-scale graphs where exact computation is infeasible [4].

8 Worked Examples

Consider a source connected to two nodes a and b via paths of equal length but differing interaction quality. Let both paths have total cost 1, but let $q = 1$ along the path to a and $q = 0.5$ along the path to b . Then $A(a) = \frac{1}{2}$ while $A(b) = \frac{0.5}{2} = 0.25$, despite identical shortest-path distance.

If a is further connected to a dense cluster of high-quality interactions within N_r while b is not, the second term in $BC(\cdot)$ further separates their scores. This illustrates that Buscemi centrality distinguishes nodes that are indistinguishable under shortest-path distance or harmonic centrality.

³For completeness, the authors have also computed the Bacon number of Paul Erdős and the Erdős number of Kevin Bacon. These quantities are finite. We do not know what to do with this information.

8.1 John Goodman

Let v be John Goodman. In a film coappearance graph, Goodman has a direct connection to Buscemi through *Monsters, Inc.*, in which Goodman voices Sulley and Buscemi voices Randall [3].

If the **coappearance** edge type is assigned traversal cost $c(\text{coappearance}) = 1$ and quality factor $q(\text{coappearance}) = 0.9$, then the best path from Goodman to Buscemi is the direct one-edge path. Its quality is

$$Q(P) = \frac{0.9}{1+1} = 0.45,$$

so $A(\text{Goodman}) = 0.45$.

Suppose the neighbourhood N_r contains a small set of Buscemi-adjacent actors, including Goodman, with weights

$$\alpha(\text{Goodman}) = 0.30, \quad \alpha(u_1) = 0.40, \quad \alpha(u_2) = 0.30,$$

and that Goodman has accessibility $A(\text{Goodman}, u_1) = 0.20$ and $A(\text{Goodman}, u_2) = 0.10$.

Then for $\lambda = 0.5$,⁴

$$\begin{aligned} BC(\text{Goodman}) &= 0.5 A(\text{Goodman}) + 0.5 \sum_{u \in N_r} \alpha(u) A(\text{Goodman}, u) \\ &= 0.5(0.45) + 0.5(0.30 \cdot 1 + 0.40 \cdot 0.20 + 0.30 \cdot 0.10) \\ &= 0.225 + 0.5(0.41) \\ &= 0.43. \end{aligned}$$

Thus $BC(\text{Goodman}) = 0.43$, reflecting both his direct coappearance with Buscemi and his embedding within the surrounding Buscemi-adjacent neighbourhood. This shows that Buscemi centrality is not merely a shortest-path quantity; vertices receive credit both for direct accessibility and for structural embeddedness within the Buscemi-adjacent neighbourhood.

8.2 Adam Sandler

Adam Sandler's coappearance relationship with Buscemi is considerably more extensive than Goodman's, spanning multiple productions including *Billy Madison* (1995) and *Mr. Deeds* (2002). Repeated coappearance reflects a more robust evidential association; we accordingly assign a distinct edge

⁴The case $\lambda = 0.5$ was selected because it produces a round number. Science.

type `repeated_coappearance` with $q(\text{repeated_coappearance}) = 0.95$ and $c(\text{repeated_coappearance}) = 1$.

The best path from Sandler to Buscemi is the direct one-edge path with quality

$$Q(P) = \frac{0.95}{1+1} = 0.475,$$

giving $A(\text{Sandler}) = 0.475 > A(\text{Goodman}) = 0.45$. With comparable neighbourhood embedding, we obtain $BC(\text{Sandler}) \approx 0.49$.

This result is not considered a deficiency of the measure.

8.3 Sample Buscemi Centralities

Table 1 presents illustrative Buscemi centrality scores for a selection of nodes. Quality factors and neighbourhood embeddings are estimated for illustrative purposes and should not be used for load-bearing applications.

Node	Best path type	$A(v)$	$BC(v)$
Steve Buscemi	(source)	1.000	1.000
Adam Sandler	direct, repeated	0.475	0.49
John Goodman	direct	0.450	0.43
Harvey Keitel	direct	0.450	0.41
Kevin Bacon	direct (<i>Sleepers</i> , 1996)	0.450	0.42
Paul Erdős	not established [†]	0.000	0.000

Table 1: Illustrative Buscemi centrality scores for selected nodes. [†]Erdős has no known path to Steve Buscemi. In death as in life, he remains an outlier.

9 Empirical Observations

We observe that:

- Nodes with identical shortest-path distance to Buscemi can differ significantly in $BC(v)$.
- High-quality interaction paths dominate over shorter but weaker paths.
- Nodes embedded in dense Buscemi-adjacent regions exhibit elevated centrality.

10 Limitations

The proposed measure has the following limitations:

- The measure is undefined for graphs containing no Steve Buscemi.
- Extension to directed graphs is left as an exercise.
- The authors have not computed their own Buscemi centrality, but assume it is non-zero.

11 Conclusion

We have introduced Buscemi centrality, a source-relative centrality measure for heterogeneous affiliation graphs. By combining best-path accessibility with source-neighbourhood context, the measure captures interaction quality and structural embedding in a way not reflected by classical distance-based metrics. We have further demonstrated that Buscemi centrality strictly generalises the Bacon number and, transitively, the Erdős number, confirming that all of graph theory can, at some level, be entirely about Steve Buscemi.

Acknowledgements

The authors thank Steve Buscemi for his extensive contributions to the source neighbourhood of the film coappearance graph, without which this measure would be considerably less interesting. We also thank the anonymous reviewers, who we assume exist.

Conflict of Interest

The authors declare no conflict of interest. Steve Buscemi was not consulted during the preparation of this work and is under no obligation to acknowledge its existence.

References

- [1] Paolo Boldi, Marco Rosa, Massimo Santini, and Sebastiano Vigna. Hyperan: Approximating the neighbourhood function of very large graphs

on a budget. In *Proceedings of the 20th International World Wide Web Conference (WWW)*, 2011.

- [2] Paolo Boldi and Sebastiano Vigna. Axioms for centrality. *Internet Mathematics*, 10(3-4):222–262, 2014.
- [3] Monsters, inc. <https://www.imdb.com/title/tt0198781/>. Accessed 24 March 2026.
- [4] Mark Newman. *Networks: An Introduction*. Oxford University Press, 2010.