Accepted to be published in: 2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC 2024), 11-13 June 2024, Chicago-USA.

Cite as:

C. Schneider and P. Szczurek, "Stability of Role Assignments in Modified Networks," 2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC), Chicago, IL, USA, 2024, pp. 1-6, doi: 10.1109/INTCEC61833.2024.10603056.
keywords: {Knowledge engineering;Corporate acquisitions;Perturbation methods;Airports;Stability analysis;Airline industry;network analysis;role discovery;RolX;dynamic networks;U.S. Department of Transportation;passenger airline network},

BibTeX:

@INPROCEEDINGS{10603056,
author={Schneider, Christopher and Szczurek, Piotr},
booktitle={2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC)},
title={Stability of Role Assignments in Modified Networks},
year={2024},
volume={},
number={},
pages={1-6},
keywords={Knowledge engineering;Corporate acquisitions;Perturbation methods;Airports;Stability analysis;Airline industry;network analysis;role discovery;RolX;dynamic networks;U.S. Department of Transportation;passenger airline network},
doi={10.1109/INTCEC61833.2024.10603056}}

# Stability of Role Assignments in Modified Networks

Christopher Schneider
Department of Engineering, Computing, and Mathematical Sciences
Lewis University
Romeoville, IL, USA
christopherjschnei@lewisu.edu

Piotr Szczurek
Department of Engineering, Computing, and Mathematical Sciences
Lewis University
Romeoville, IL, USA
szczurpi@lewisu.edu

***Abstract*—This work introduces the concept of the role-stability coefficient, which can be used for measurement and interpretation of unsupervised role assignments. We also present the idea of a role-shift matrix and show how these concepts can be used for studying impacts caused by instantaneous and significantly scaled perturbations to the existing network, such as by appending a significantly sized second network to the first. We describe and provide an example of an application to the airline industry and show how our method can be used to provide valuable in-depth insights into how airline mergers can impact the roles of airports.**

***Keywords*—network analysis, role discovery, RolX, dynamic networks, U.S. Department of Transportation, passenger airline network.**

## I. INTRODUCTION

While the nodes of networks are in large part defined by their connections to other nodes, there are many contexts in which the node itself can be thought to possess or create forces that exert influence on the broader network. For applications focused on the identity of nodes themselves, the concept of role discovery is an essential tool. Role discovery is the process by which we can label network nodes based on their structural or functional similarity [1]. Applications for role discovery extend to dynamic networks ([2], [3]). Much of the work tends to be focused on organically propagating changes and small discrete time-step sizes, essentially trying to model user/node activity in something approximating real time. In this work, we wish to provide an essential tool to study a different sort of semi-dynamic network. The scenario of interest is one in which we may not be concerned with a temporal element, as the change applied to the network comes instantaneously and with more substantial relative scale. As a starting point, we will consider the scenario of appending a second network, fully formed, onto the first. In doing so, the two networks become one, and the characteristics of each may be expected to fluctuate when subject to perturbations posed by the other. This phenomenon may manifest itself by changing the role assignment for certain nodes.

We propose a method for quantifying and interpreting this role-shifting behavior. Our objective was to anticipate the impact of change to an existing structure at a topological level. We focus our attention on intuitively network-mappable systems in which the roles of nodes are not merely labels to be assigned but signify vital functions, and where that function may be fluidly assigned rather than positionally fixed. These underlying functions may best be served in different ways—changes in control structure, logistics, and other inputs/outputs tend to evolve over time.

By understanding the propensity of role-shifting within the network, we aim to quantify and interpret those changes into measurable outcomes. Networks that encounter a lower degree of perturbation can be expected, on balance, to prove more resilient and experience less operational friction or transition cost to implement the change. In short, we seek to apply a dimensional measurement to the nebulous idea of "synergy."

There are many real-world scenarios that closely fit the scheme described. Social media companies may see applications tied to acquiring and integrating a new user base with its own community structures and relationships already in place. Transit network planners may wish to understand the impact of building a link between routes or a hub connecting multi-modal services. Opening that link or hub has the effect of joining two networks, with all the potential for impacts as described above. In another example, a company which operates numerous outlets may seek to change its footprint in some way, such as a shopping center redesigning traffic corridors or installing a new anchor, or a fast-food chain franchising additional stores within a partially saturated neighborhood. In general, operations analysis will find a natural fit with assessing merger and acquisition proposals.

In this work, we will focus on studying role-shift in a airline network analysis. In our scenario, major airline networks seek expansion opportunities by extending alliances to an array of smaller regional carriers. We will show how our proposed method allows us to determine which airline acquisitions may be implemented with the least perturbation to the merged network.

The rest of the paper is structured as follows. In section II, we will provide some necessary definitions and formally state the problem of role stability. Section III will outline the methodology for our proposed solution. In section IV, we will describe a potential application and how we applied our method in a hypothetical scenario. This is followed by our conclusions in Section V. Related work is described in section VI.

## II. DEFINITIONS

A network is a weighted graph $G$, which is defined as a set $(V, E, w)$, where $V$ is a set of vertices, $E$ is a set of edges that connect pairs of vertices, and $w: E \rightarrow \mathcal{R}$ is a function that maps edges to real numbers. Given two weighted graphs, $G_1 = (V_1, E_1, w_1)$ and $G_2 = (V_2, E_2, w_2)$, a *merged graph* $G_m = (V_m, E_m, w_m)$, where $V_m = V_1 \cup V_2$, $E_m = E_1 \cup E_2$, and $w_m = w_1 + w_2$, that is, the weights of the edges in the merged graph are the sums of the weights of the same edges in the constituent graphs.

Given a graph $G = (V, E)$, a *role assignment* is a mapping $ra: V \rightarrow R_G$, that maps each vertex V of G to a finite set of values, called *roles* from the set $R_G$. This assignment is generated using a *role discovery algorithm*. The number of roles for a graph G, $|R_G|$, may be different for distinct graphs. The set of all vertices from graph $G$ that have the same role $r$ is called a *role set*, and is denoted by $\{r_G\}$. We denote an ordered list of all role sets for a given graph as $[\{r_G\}]$, which we call the *role list*. For two graphs $G_1$ and $G_2$ and two sets of roles, $R_{G1}$ and $R_{G2}$, derived by some role assignment, we define the *role matching problem* as finding a mapping between $R_{G1}$ and $R_{G2}$ that maximizes a specified similarity metric. This metric may depend on the specific data to which the problem is applied, but we will present a general method that can be used in section IV. For convenience, we will assume the similarity metric values range from zero to one, where a value of one indicates a perfect similarity. This means the mapping matched identical roles and all roles were matched.

Let $G_s$ be a single source graph, which we want to merge with one of $N$ target graphs $G_{t1}, G_{t2}, \ldots, G_{tN}$. After mering the source graph $G_s$ with some target graph $G_{ti}$, we can then get a new merged graph $G_{s+ti}$. Assume we perform role discovery to find a role assignment and determine the optimal role matches. If the role matching generates a perfect similarity, we say it exhibits *role stability*. A *role stability measure* quantifies the degree to which the role matching was not perfect.

## III. Role-Stability Measurement

In this section, we describe our proposed method for deriving a role stability measure we call the *role-stability coefficient* for a given source graph and another target graph. The method breaks down into several distinct phases, which are shown in Algorithm 1. First, we merge the source graph with the target to get a merged graph, combining identical nodes and overlapping edges as needed. Second, we perform unsupervised role discovery on the source network, our object of study, and the merged network. Third, we perform role matching between the source and the merged network. This generates the matches and the similarity values of the matched pairs. We then compute the role-stability coefficient as the sum of the similarities of matched pairs, normalized by the number of roles in the source graph. This coefficient is a real value in the range of $[0, 1]$, and can be used to directly compare against like measurements from other potential network augmentations. Larger values provide the greater amount of role stability, with a value of one representing a perfect alignment with zero shifting effect.

The algorithm we outlined relies on specifying two procedures: performing role discovery and role matching. We will discuss potential choices for role discovery in the next subsection. We will then explain our proposed approach for role matching in subsection B. Lastly, in subsection C, we propose a way to gain further insights into results with a role-shift matrix.

**Algorithm 1** role_stability_coeff($G_s, G_t, RD$)
Computes the Role-Stability Coefficient

Inputs: source graph $G_s = (V_s, E_s)$, target graph $G_t = (V_t, E_t)$, role discovery algorithm RD($G$) which returns a role assignment function

```
G_{s+t} ← merge(G_s, G_t)  // get merged graph
ra_s ← RD(G_s)  // get role mapping for source graph
ra_{s+t} ← RD(G_{s+t})  // get role mapping for merged graph
R_s ← ra_s(V_s)  // get roles for source graph vertices
R_{s+t} ← ra_{s+t}(V_{s+t})  // get roles for merged graph vertices
// find matches between roles of source and merged graphs, along with the corresponding similarities
    RM = [(r_s, r_{s+t}, sim)] ← role_matching([{r_s}], [{r_{s+t}}])
// sum the similarity scores for each match
    match_score ← Σ_i RM_i(sim)
return role_stable_coeff ← match_score / |R_s|
```

### A. Role Discovery

Conducting role discovery is the cornerstone of this approach, so it makes sense to be thoughtful in selecting an algorithm for this. While we are very much interested in the functions that nodes play, the shifting dynamics we aim to study belie the implicit supposition that these functional roles may be fluid. We wish for the role discovery process to make the optimal shifts from a topological perspective, relying on this assumption to allow functional roles to shift in turn.

We chose the RolX algorithm for role discovery due to a number of characteristics suitable for our purpose, although our method is not limited to its use [4]. As specified, RolX conducts topological analysis rather than functional, focusing on matrix reductions to optimize a Frobenius norm computation on modularity. It is an unsupervised method, so virtually no pre-analysis preparation beyond data formatting need be conducted to use it. It runs without requiring hyperparameter inputs, removing a potential source of inconsistency or user error. It is flexible, taking into account edge weighting where present while not requiring it.

The selection of RolX also poses some challenges which we need to address. Chiefly, the unsupervised nature of the algorithm leads to behavior which is difficult to interpret directly. For example, in generating the illustrative example in section IV, we encountered instances in which the RolX algorithm's output produced gaps in sequential role naming, an artifact of roles being generated as the algorithm proceeds and then eliminated during soft clustering resolution—and an inscrutable result for an end user to encounter.

As the algorithm runs without prior knowledge, the identities of roles are not implicitly preserved and may not be assessed identically on non-identical graphs, making the direct measurement of role shifts a non-trivial task. Therefore, the roles

**Algorithm 2** role_matching ($[\{r_1\}], [\{r_2\}]$)
Finds the optimal matching of roles from two graphs

Inputs: role lists for two graphs, $[\{r_1\}], [\{r_2\}]$

**for each** role set $\{r_1\}$ **in** $[\{r_1\}]$
  **for each** role set $\{r_2\}$ **in** $[\{r_2\}]$
    $\{r_2\}' \leftarrow \{v: v \in \{r_1\}\}$
    $sim\_arr[r1, r2] \leftarrow |\{r_1\} \cap \{r_2\}'| / |\{r_1\} \cup \{r_2\}'|$
$[(r_1, r_2, sim)] \leftarrow lin\_maxsum\_assign(sim\_arr)$
**return** $[(r_1, r_2, sim)]$

between two graphs have to be matched in attempt to compare them. This is described in the next subsection.

### *B. Role Matching*

The problem is to match roles discovered from one graph to roles discovered in another graph. Our proposed method is shown in Algorithm 2, above. We perform the role matching process by modeling the problem as a linear sum assignment. The two sets of roles, found by applying the role discovery algorithm to the two graphs, can be viewed as forming a weighted bipartite graph, with edges connecting all pairs of vertices, which we will refer to as the *role-matching graph*. The edges of this role-matching graph are weighted by the Jaccard similarities between sets of nodes in each respective role. This defined as the ratio of the size of the intersection of sets, divided by the size of the their union. We only consider nodes from the merged graph that also appear in the source graph. This is done to eliminate the influence of new nodes added from the target graph. Then, problem is to find the set of edges that maximize the sum of the weights. The edges have to be chosen such that the mapping is one-to-one. That is, each role from the first graph is matched to only one role in the second graph, and vice-versa. We find the optimal matching using the Jonker-Volgenant algorithm, which can be used on unbalanced problems, where there might be more roles in one graph than another [5]. The whole role matching process is outlined in Algorithm 2. Note that the Jonker-Volgenant algorithm is typically executed to minimize the sum of matched edges, but we present the problem here as an equiavelent maximization problem. The implementation we used was through python's scipy package, which allows maximization to be specified as a setting [6].

As an example of how the algorithm is applied, assume we have a source graph with vertices $V_s = \{A, B, C, D, E\}$ and a target graph with vertices $V_t = \{A, B, C, E, F\}$. The merged graph will then contain vertices $V_{s+t} = \{A, B, C, E, D, F\}$. We then apply role discovery to both source and merged graphs. Let the role list for the source graph be $RL_s = [\{A, B\}, \{C\}, \{D, E\}]$, and the role list for the merged graph be $RL_{s+t} = [\{A, B, C\}, \{D, E, F\}]$. We then compute the Jaccard similarities for all pairs in $RL_s \times RL_{s+t}$, but eliminating vertex $F$ from the second set of $RL_{s+t}$. We can represent all similarities as a two row by three column matrix, that is then used as input to the Jonker-Volgenant algorithm. The results are shown in fig. 1. From this resulting role match, we can then compute the role-stability coefficient, which is $(2/2 + 2/3)/3 = 5/9 \approx 0.56$.

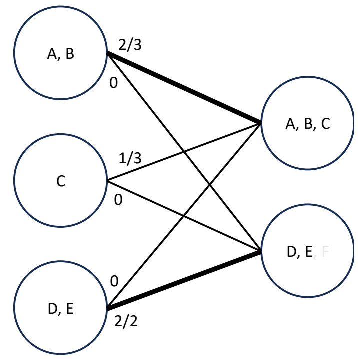


Fig. 1. Example of a role match bipartite graph, with Jaccard similarities computed between sets of roles, and the optimal assignment shown.

This shows a rather significant shift, which emerged due to the compression of roles in the merged graphs.

### *C. Role-Shift Matrix*

While the role-stability coefficient fulfills our stated objective to quantify the extent of perturbation caused by external impact to a network, it is rather opaque when it comes to understanding the underlying nature of the change which has taken place. To gain further understand, we propose to compose a *role-shift matrix*, $RSM([\{r_1\}], [\{r_2\}])$, with $|R_1|$ rows and $|R_2|$ columns, representing the $|R_1|$ unique roles identified within the source network and the $|R_2|$ corresponding roles in the merged network. Entries represent counts of nodes, with the $n_{i,j}$ entry corresponding to the number of nodes assigned to role $i$ initially and role $j$ subsequent to the impact. With the role-shift matrix, and with each role's characteristics determined by aggregating or averaging the attributes of nodes contained within, it becomes possible to answer a diverse array of functionally significant. The closer the role-shift matrix is to a square, diagonal matrix, the less role-shift can be said to have occurred. An example of the role-shift matrix for the previous example is shown below.

$$RSM([\{r_1\}], [\{r_2\}]) = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 3 \end{bmatrix}$$

We add to note that, with some flexibility to our point of reference, we can broaden the potential scope of this method. It is possible to evaluate role-matching scenarios for shrinking networks, which lends itself to networks similar to the applications given above but where operators are seeking to reduce or partition footprint. Similarly, we can assess network impacts in which the nodes themselves are static but edges or edge weighting are modified.

## IV. Application: U.S. Airline Network Analysis

To demonstrate the utility and efficacy of our proposed role-stability analysis approach, we apply our method to the domain of the airline industry. We consider a hypothetical scenario, in which one large airline or an alliance or several airlines, would like to acquire another in the hopes of merging their flight networks. While such decisions are made through consideration of a multitude of factors, airlines will likely want the acquisition process to limit the disturbance to their existing infrastructure. Therefore, we assume the roles played by the airports utilized by the airline should remain relatively constant. Role-stability analysis can provide information as to which acquisition target airline would cause the least disruption.

We use data from the TranStats portal maintained by the Bureau of Transportation Statistics, an office under the U.S. Department of Transportation. The specific reference data is the "T-100 Domestic Market (All Carriers)" dataset for the period of January through August 2023 [7]. This dataset contains aggregated counts for passengers, freight and mail cargo, carrier (airline) identification tags, and airport FAA location identifiers (three-letter codes, such as "ATL" for Hartsfield-Jackson Atlanta International Airport) for the flight route origin and destination airports. Service between pairs of airports is listed with separate records for each direction.

Using this data, we constructed a weighted digraph with the nodes representing airports, and a directed edge representing the total flow of traffic from the origin to the destination. When multiple carriers offer service for the same route, we treat these as separate edges for the sake of cleanly isolating a carrier's initial network state, with each edge associated with the number of passengers transported by that airline. We identify four major airlines/alliances for our analysis, which are shown in Table I, below [8].

TABLE I. MAJOR AIRLINES CHOSEN FOR CONSIDERATION.

| Airline / Alliance | Constituent Airlines | Number of Routes |
|---|---|---|
| Oneworld | American Airlines Inc.<br>Alaska Airlines Inc.<br>Envoy Air<br>PSA Airlines Inc.<br>Hawaiian Airlines Inc.<br>Horizon Air<br>Piedmont Airline | 2,673 |
| Star Alliance | United Air Lines Inc.<br>CommuteAir LLC (DBA CommuteAir)<br>Air Wisconsin Airlines Corp<br>GoJet Airlines LLC (DBA United Express) | 1,298 |
| SkyTeam | Delta Air Lines Inc.<br>Endeavor Air Inc. | 1,293 |
| Southwest Airlines Co. | No alliance partnerships, but their standalone scale is comparable to the 3 networks above. | 1,582 |

Note: several alliances contain many more partners than shown, but as our study is limited to U.S. domestic air travel, major international carriers such as KLM or Air Canada were generally not in scope for this illustration.

We then select six minor airlines as potential alliance or acquisition targets. These were not, based on public-facing marketing information, members of any major alliance or partnership agreement at the time of research.

- Mesa Airlines Inc. (309 route edges)
- Sun Country Airlines d/b/a MN Airlines (164)
- Breeze Aviation Group (162)
- TEM Enterprises d/b/a Avelo Airlines (142)
- Delux Public Charter LLC (31)
- Silver Airways (23)

We conducted data cleaning using common and general knowledge of the airline industry. For example, we consolidated airport codes where the airport service moved or changed during the span of service; such outliers were identified by seeking airport codes with latitude/longitude placing them in near proximity, then doing cursory research to validate what we found in the dataset. We removed any flight routes for which an airline did not offer a reciprocal route in the opposite direction, as these were all zero-passenger or low-passenger records and seemed most likely to represent unscheduled or emergency landings, such as redirects due to inclement weather.

We prepared networks representing the source graph for the four major alliances, as well as 24 graphs representing the synthesized hypothetical networks formed by the Cartesian product of merging each combination of one major and one minor airline. Where both airlines in a synthesis offer the same route, we added the passenger counts to weight a single directed edge representing the net traffic from both airlines.

Once the array of potential networks was created, we performed role discovery using the RolX algorithm on each of the 24 synthesized networks, along with the four original major alliance networks as baseline information, seeking to classify the roles of all airport nodes in each network. For example, in the analysis of SkyTeam acquiring Mesa Airlines in its alliance, the role classifications of SkyTeam alone were compared to the same nodes in SkyTeam+Mesa.

We ran the process described in section III, including construction of role-shift matrices. Tools and libraries used were Python 3.11.4 with standard packages via Anaconda [9]; Spyder and Jupyter Notebook programming environments; pandas, numpy, scipy, matplotlib including seaborn and basemap, NetworkX, and GraphRole external libraries [10], along with Microsoft Excel from Office 365 to perform last-mile data manipulation to gather statistics and prepare role-shift matrices via PivotTable.

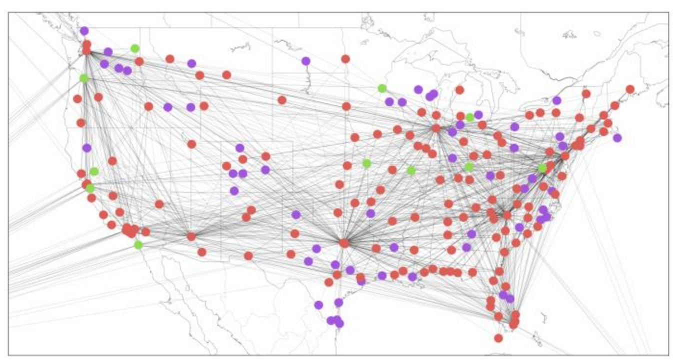

Fig. 2. Roles for the original network for the Oneworld alliance.

To assess the prima facie interpretation of the role-stability coefficients, we prepared maps showing the network nodes and edges both before and after. Figure 3 shows the original network with different roles corresponding to vertex colors. In fig. 3, we show the roles after merging the original Oneworld alliance network with Silver Airways. Visual comparison to the previous figure shows the roles were largely preserved. In fig. 4, we show roles of OneWorld after merging with Delux Charter, which

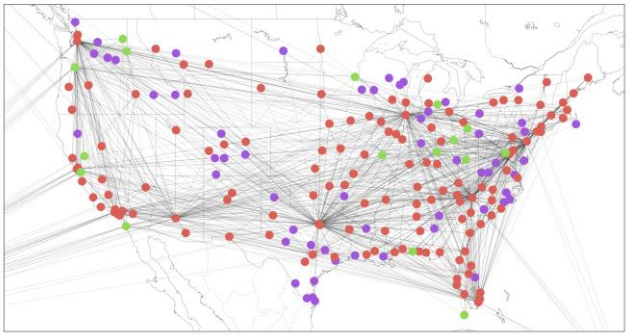

Fig. 3. Roles with a merged Oneworld alliance and Silver Airways, which produce the least role shift.

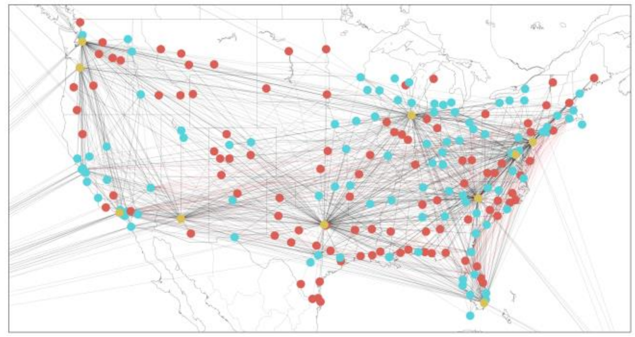

Fig. 4. Roles for the merged Oneworld alliance and Delux Charter, which generated the largest role shift.

TABLE I. ROLE-STABILITY COEFFICIENTS FOR ALL COMBINATIONS OF MAJOR AND MINOR AIRLINE NETWORKS

| | **Original** | **Mesa Airlines** | **Sun Country** | **TEM-Avelo** | **Breeze Aviation** | **Delux Charter** | **Silver Airways** |
|---|---|---|---|---|---|---|---|
| Oneworld | 1 | 0.28 | 0.39 | 0.49 | 0.28 | 0.23 | **0.87** |
| Southwest | 1 | 0.17 | 0.24 | **0.46** | 0.31 | 0.24 | 0.28 |
| Star Alliance | 1 | 0.18 | 0.16 | 0.54 | **0.70** | 0.26 | 0.44 |
| SkyTeam | 1 | 0.24 | 0.27 | 0.22 | 0.24 | 0.33 | **0.54** |

ended up producing the largest role changes. After such visual inspection is helpful, by computing the role-stability coefficient and role-shift matrices, we can get quantitative knowledge about the extent of the role changes and the nature of those changes. The role-stability coefficients for all combinations of major and minor airlines are shown in Table II, below.

The results match the analysis from the visual inspection, indicating the merger of Oneworld with Silver Airways would produce the greatest role-stability, while merging with Delux Charter would generate the least role-stability. In some cases, the use of alternate airports in the same market created complications in likely passenger routing rather than streamlining. This was frequently true with Southwest Airlines, which has historically used second-tier airports for major cities such as Dallas Love Field or Chicago Midway as an operating strategy, and which had the lowest role-stability coefficient of the four networks for two of the six proposed partnerships.

To gain further insights into how roles are shifted for different merged networks, we used role-shift matrices to attempt to provide narrative description of the role shifts that had taken place in several different scenarios. A highly diagonal role-shift matrix corresponded to a network pairing with high role stability and thus a higher role-stability coefficient. In addition, we had computed several centrality measures for each node of the network, to better explain the results provided by the the role-shift matrix. These included betweenness, PageRank, hubs and authorities, closeness, and eigenvector centrality (see [11] for explanations of the measures).

### *A. OneWorld + Sun Country*

As shown in fig. 5, the transition matrix is somewhat diagonal, indicating overall role stability through the network modifications. We see that of 18 cities originally in role_0 (the largest nationwide hubs including DFW, ORD, JFK, and LAX) within Oneworld's alliance network, 13 remained in that same role, 6 new cities were added and 5 shifted to different roles, for

| **Oneworld+Sun Country** | **x-axis: transition class** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **y-axis: original class** | **role_0** | **role_1** | **role_2** | **role_4** | **role_5** | **role_6** | **role_7** | **Grand Total** |
| role_0 | 13 | | | | 3 | 1 | 1 | 18 |
| role_1 | 1 | 2 | | | | | | 3 |
| role_2 | | | 105 | 12 | | | | 117 |
| role_4 | | | 1 | 6 | 4 | 1 | 12 | 24 |
| role_5 | 4 | | 2 | | 1 | | | 7 |
| role_6 | 1 | | 1 | 4 | | 4 | 1 | 11 |
| role_7 | | | 33 | 3 | | | 21 | 57 |
| **Grand Total** | **19** | **2** | **142** | **25** | **8** | **6** | **35** | **237** |

Fig. 5. Role shifts for the merged Oneworld alliance and Sun Country.

an updated total of 19 role_0 nodes. We see an apparent transition cycle, with 33 nodes shifting from role_7 to role_2 (smaller cities like College Station, TX, South Bend, IN, and Erie, PA); role_2 has 12 airports shifting to role_4 (Fresno, CA, Denver, CO, Columbus, OH); and role_4 sees 12 shifting to role_7 (all edge nodes in Alaska or Pacific NW). This actually can be seen with greater clarity via centrality statistics: the 7-to-2 shift reflects a slightly larger-than-norm net decrease in closeness and eigenvalue centrality with no net change in betweenness (edge nodes shifting farther from the center of mass). The 2-4 shift is characterized by positive shifts in betweenness with a few exceptions; and the 4-7 shift sees a group of similar-distance edge nodes gaining some relative connectedness, as decreases in hubness and authority are below the overall trend. With a network expansion, it is logical to expect a decentralization impact in aggregate, so any node that increases in centrality is an interesting anomaly showing unique value within the merged network.

The dropped hubs were TPA (Tampa), BNA (Nashville), RDU (Raleigh-Durham), OGG (Kahilua, HI), and GST (Gustavus, AK)--Hawaii and Alaska were excluded from the map to preserve data density. While nearly all role_0 airports declined in hubness, these declined less on average (-2.09e05 vs. -5.75e05) and RDU actually increased, and betweenness scores rose for 4 and stayed flat for 1, against a trend of decline for role_0 generally. This makes sense as the routes added tend to center around these areas, adding 1st and 2nd-degree connections that can offer more routing options for these nearby cities. The added hubs were Anchorage, Atlanta, Orlando, Pittsburgh, Portland (OR), and Seattle. These nodes play more prominent roles in other alliances compared with Oneworld's starting state, and the expanded importance of these airports represents a synergistic opportunity to compete. The data bears this out with above-trend shifts in hubness and betweenness scores, with Orlando being the single largest increase in the entire network.

### *B. SkyTeam + Silver Airways*

| Count of Silver Airways | x-axis: transition class | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| y-axis: original class | role_1 | role_2 | role_3 | role_4 | role_5 | role_6 | role_7 | Grand Total |
| role_1 | 57 | 3 | 5 | | | 1 | 1 | 67 |
| role_2 | 7 | 34 | 12 | | 1 | | | 54 |
| role_3 | | 8 | 36 | | | | | 44 |
| role_4 | | | | 3 | | | | 3 |
| role_5 | | | | | 8 | | | 8 |
| Grand Total | 64 | 45 | 53 | 3 | 9 | 1 | 1 | 176 |

Fig. 6. Roles for the merged SWA alliance and Delux Charter, which generated the largest role shift.

The role-shift matrix for SkyTeam's alliance merger with Silver Airways is shown in fig. 6. Delta Air Lines, the preeminent partner in the SkyTeam Alliance, is based out of Atlanta, and it operates a very focused hub-spoke model (once again exemplified by a narrow range of distinct role classes in starting state). Thus, many of its regional flights to Florida and the Gulf Coast are served directly out of Atlanta, which has a betweenness score of 0.58–the highest of any network we analyzed. This proposed alliance points to the efficiency gains possible by allowing more neighborhood-density to form in this tourist cluster.

In this network, the largest hubs of ATL, MSP (Minneapolis), and DTW (Detroit) all fall into role_4, which shows role stability. Role_5 is likewise stable, consisting of major destinations such as New York (JFK and LGA), LAX, SEA, and BOS. SLC, a smaller hub for the Mountain West region, also falls into role_5, and Fort Lauderdale ascends from role_2 to role_5 as the only major hub shift seen here. Other nodes gaining traction, with increases in all centrality metrics, include Orlando, Pensacola, and Key West.

Airports shifting between role_2 and role_3 are rather common. Role_3 generally represents cities that are not edge nodes, but still exhibit less network traffic and low betweenness scores. Role_2 airports tend to be edge nodes with higher traffic; all have betweenness of zero, along with role_1. 2-3 shifts include Buffalo, Honolulu, Houston-IAH, Indianapolis, Chicago-ORD, and Portland; several larger cities that are not central locations served by Delta or associated airlines.

## V. Conclusions

The major contribution of our work is to introduce the concept of role stability and the proposed role-stability coefficient measure, which quantifies the role shift dynamics of modified networks. Additionally, we propose to use role-shift matrices to study how changes break-down for a particular scenario. We show the method can be used to gain in-depth knowledge about possible network modifications by applying the techniques to the airline industry.

## VI. Related Work

Role discovery has been essential to recent advances in data mining methods ([1], [11]). Most of the work on role discovery focuses on using either supervised or unsupervised approaches to embed network nodes in a way that allows for grouping them into common structural or functional roles [12]. The novelty of our work lies in the analysis of the changes in the roles when networks are modified. While there are existing approaches for quantifying graph structure changes ([13]), they do not take into account the concept of a role. The most directly relevant work has been that in the analysis of roles for dynamic networks ([2], [14]), but this largely studying incremental changes over short temporal spans, such as social media network activity. Here, we examine change at greater scale and all at once, augmenting the existing network by appending a significantly-sized graph.

## References


[1] R. Rossi and N. Ahmed. "Role Discovery in Networks." *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 04, pp. 1112-1131, 2015.

[2] R. Rossi, B. Gallagher, J. Neville, & K. Henderson. "Modeling Dynamic Behavior in Large Evolving Graphs," in *Proc. Sixth ACM International Conference on Web Search and Data Mining* (WSDM), 2013, pp. 667-676.

[3] G.H. Nguyen, J. B. Lee, R. A. Rossi, N. K. Ahmed, E. Koh, & S. Kim.. "Dynamic network embeddings: From random walks to temporal random walks," in *2018 IEEE Int. Conf. on Big Data*, pp. 1085-1092.

[4] K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu, L. Akoglu, D. Koutra, C. Faloutsos, & L. Li. (2012). "RolX: Structural role extraction & mining in large graphs," in *KDD'12 - 18th Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2012, pp. 1231-1239.

[5] R. Jonker, A. Volgenant. "A shortest augmenting path algorithm for dense and sparse linear assignment problems." *Computing* vol. 38, pp. 325–340. 1987.

[6] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, & P. Van Mulbregt. "SciPy 1.0: fundamental algorithms for scientific computing in Python." *Nature methods*, *17*(3), pp. 261-272. 2020.

[7] *Bureau of Transportation Statistics - T-100 Domestic Market (All Carriers)*. U.S. Department of Transportation. [Online.] Available: https://www.transtats.bts.gov/Tables.asp

[8] "List of airlines of the United States." Wikipedia. [Online.] Available: https://en.wikipedia.org/wiki/List_of_airlines_of_the_United_States

[9] *Anaconda Software Distribution v.2-2.4.0*. (2016). Anaconda. Accessed: Oct. 23, 2023. [Online.] Available: https://anaconda.com

[10] *GraphRole*. D. Kaslovsky. Accessed: Dec. 6, 2023. [Online.] Available: https://github.com/dkaslovsky/GraphRole

[11] P. Szczurek, & M. Horeni. "Using Link Analysis Algorithms to Study the Role of Neurons in the Worm Connectome." In *2018 IEEE 32nd Int. Conf. on Adv. Info. Networking and Applications (AINA)*, pp. 651-657.

[12] P. Jiao, X. Guo, T. Pan, W. Zhang, Y. Pei, & L. Pan. "A survey on role-oriented network embedding." *IEEE Trans. Big Data*, vol. 8, no. 4, pp. 933-952, 2021.

[13] P. Wills, F. G. Meyer "Metrics for graph comparison: A practitioner's guide." *PLoS ONE*, vol. 15, no. 2, e0228728. doi: 10.1371/journal.pone.0228728

[14] R. Rossi, B. Gallagher, J. Neville, & K. Henderson. "Role-dynamics: fast mining of large dynamic networks." In *Proc. 21st Int. Conf. on World Wide Web*, Apr. 2012, pp. 997-1006.