

# A Synthetic Model for Multilevel Air Transportation Networks



Marzena Fügenschuh, Raluca Gera and Tobias Lory

## 1 Motivation and Related Work

The global transportation system is a very dynamic and intricate network. Optimizing travel through this network to efficiently transport goods and people via air travel, as well as analyzing its resilience to disruption, is highly desirable. Based on the real-world limitations of airports, aircrafts, financial and personnel resources as well as the unpredictability of weather and natural disasters, many variables must be taken into account. In order to effectively study the real world development of this complex network, methodical means of creating synthetic networks comparable in scope and behavior to real world data are needed. The natural development of air transportation networks is difficult to model because of the multilayered nature of the networks. Each airline independently creates routes based on market analysis for profit, competitor routes and available resources and destinations. On the other hand, each airport is separately developed by the municipalities it services with input and oversight from national and international governing bodies.

One way that this network has been studied in the past is through the analysis of multilayered networks. Multilevel or multilayered networks, frequently referred to as multiplexes, have been considered as a detailed extension of the single layered networks [1–3]. This structure is desirable in our case, as each airline company can easily be modeled by a layer, with the airports being captured by the nodes. While generating synthetic networks [4] has been very active research area, less has been done in synthetic multilayered network generation [3]. In the most common approach

---

M. Fügenschuh (✉) · T. Lory  
Beuth University of Applied Sciences, Berlin, Germany  
e-mail: fuegenschuh@beuth-hochschule.de

R. Gera  
Department of Applied Mathematics, Naval Postgraduate School, Monterey,  
CA 93943, USA

growing multiplex network models are based on preferential attachment [5, 6] as they usually model relations in social networks.

Particular attention has been paid to the European Air Transportation Network (EATN), studied in [7]. A model for the network was introduced in [8], where the scale-free structure of airline networks is exploited and models simulating air traffic network based on preferential attachment are introduced. However, these models do not exploit the multilayered structure. In [9] the multilayer and the scale-free structure of EATN is exploited to design a generative model based on an enhanced preferential attachment method to imitate the EATN. As investigations of existing air transportation networks confirmed their scale-free nature [8], the approach of Barabási-Albert comes in handy to model the layers of this network. The preferential attachment method can indeed deliver a reliable multiplex network model [9]. However, the inter- and intra-layer structure has not been considered in detail.

In the current work, we build on the *BinBall* model using the Barabási-Albert approach to model the diversity of the layers within a multiplex network.

## 2 An Enhanced Synthetic Model for a Multiplex

A multiplex as a complex network consists of several layers (subnetworks), on the same set of nodes. As each layer is given by a different attribute (different airline in our case), the edges of the layers may duplicate each other. Thus, the multiplex  $M$ , is an undirected multigraph consisting of simple undirected graphs, the layers,  $L_1, \dots, L_\ell$ , for some  $\ell > 1$ , i.e.  $M = \bigcup_{k=1}^{\ell} L_k$ . A node of a multiplex can be viewed within a single layer, or globally in the whole network. Thus one distinguishes between the *local degree* of a node  $u$  with respect to some layer  $L$ ,  $\deg_L(u)$ , and the *global degree* with respect to the multiplex,  $\deg_M(u)$ .

In the *BinBall* model [9], an empty network on the node set shared across all layers is initialized. The node set is divided into possibly equally-sized subsets indicating the layers. Edges are added iteratively. For each edge,  $e = (u, v)$ , the layer  $L$  is chosen randomly. The selection of the end nodes is based on their local and global degrees. The probability of a node  $u$  being chosen as the first end-node of an edge, and a node  $v$  as the second end-node is:

$$\frac{\alpha \deg_L(u) + s}{\sum_{t \in V_L} (\deg_L(t) + s)} \quad \text{and} \quad \frac{\alpha \deg_M(v) + P(v) + s}{\sum_{t \in V} (\deg_M(t) + P(v) + s)},$$

respectively. Here,  $\alpha$ ,  $s$  and  $P$  are predefined values:  $\alpha$  is a scaling factor mapping a node degree to a weight,  $s$  the *zero appeal* - a base value added to all nodes' weights when randomly choosing a node, and  $P$  a mapping from the nodes to positive reals indicating a node's global weight.

The *BinBall* model simplifies the multiplex structure, because a unified evolution manner is applied to all layers. As a result, layers of similar node and edge

sizes contribute to the network. All layers evolve alike with respect to their degree distribution.

We introduce *StarGen*, a model summarized in Algorithm 1, that focuses on the diversity of the distinct layers within a multiplex. Inspired by *BinBall*'s preferential attachment we create an asynchronous growth of the layers in the multiplex. To do so, we allow different sizes of the layers based on a predefined distribution of layers' edge count. Furthermore, we decouple the scaling factor  $\alpha$  by distinguishing between *local* and *global*  $\alpha$ -values. We vary the local  $\alpha$ -values to influence the variety of the intra-layer structure: to each layer  $L_k$ ,  $1 \leq k \leq \ell$ , we assign  $\alpha_k$  as the layer's own local exponent. We consider

$$\frac{(\deg_L(u))^{\alpha_k}}{\sum_{t \in V_L} (\deg_L(t))^{\alpha_k}} \quad (1)$$

as the probability of a node  $u$  being chosen as the first end node, as well as,

$$\frac{\alpha \deg(v) + s}{\sum_{t \in V} (\alpha \deg(t) + s)}, \quad (2)$$

the probability of a node being chosen as the second end node.

---

### Algorithm 1 *StarGen*

---

#### Input

$l, m, n$  - the total number of layers, edges and nodes in the multiplex, resp.

$s$  - zero appeal,  $\alpha$  - global  $\alpha$ -value,  $\alpha_1, \dots, \alpha_l$  - local  $\alpha$ -values

$P_L^E = (p_1, \dots, p_l)$  - layer edge sizes distribution

1: initialize multiplex  $M$  on  $n$  nodes, and empty layers  $L_1, \dots, L_l$

2: **for** each edge  $e \in 1 \dots m$  **do**

3:   select a layer, say  $L_i$ , with respect to  $P_L^E$

4:   **if**  $node\_size(L_i) \leq 0.25 \cdot n$  **then**

5:     select start node  $u$  according to the local preferential attachment (1)

6:     select end node  $v$  according to the global preferential attachment (2)

7:   **else**

8:     select start and end node  $u, v$  randomly from nodes in  $L_i$

9:   **end if**

10:   add the edge  $e = (u, v)$  to layer  $L_i$  and to multiplex  $M$

11:   update local and global degree distribution of  $u$  and  $v$  according to (1) and (2)

12: **end for**

**Output**  $M, L_1, \dots, L_l$ .

---

The layer's sizes evolve via the preferential attachment. To avoid very large layers we enforce a random selection of both nodes from the layer, if its node count exceeds 25% of the multiplex node size.

### 3 Data Analysis and Model Validation

Following [9] we validate our model with a real-world multiplex network data of [7]. In airline networks, nodes represent airports and edges represent flights between two airports on a given airline. A layer in this network represents the contribution of a particular airline to the network. As already reported in [7] the EATN consists of 450 distinct node labels, 37 layers, and 3588 edges (including duplicates from different layers). The layers, especially those corresponding to national airlines, tend to build a hub and spoke structure. The emergence of a hub in one layer makes it a good candidate for a spoke in another layer. As a result, the multiplex as the union of all layers has a power law degree distribution.

Our analysis of the inner, layered structure of the network revealed that the layers vary from 35 to 128 nodes, and from 34 to 601 edges. While the layer’s sizes based on nodes are nearly uniformly distributed, the edge counts follow a power law distribution. Although almost all layers resemble hub and spoke structure, it shapes differently over the layers. We deduce it from the highly volatile percentage of one degree nodes across the layers, see the first chart on the left in Fig. 1. Each color represents the group of nodes of degree 1, followed by the ones of degree less than  $t\%$  of local maximum degree, where  $t \in \{10, 20, \dots, 100\}$ . For each  $x$ -value representing a layer, the  $y$ -value is the count of each color group, normalized by the layer’s node count.

We measure the performance of the *StarGen*-model by comparing it to the *BinBall*-model and EATN. We sample 100 synthetic networks of both models with common input values for  $\ell = 37$ ,  $m = 3588$ ,  $n = 450$ , and  $\alpha = 1.0$ . In *BinBall*-model, the  $P$ -values represent node degrees of a random preferential attachment graph on the multiplex’s node set, with incoming nodes attaching with one edge, and  $s$  is set to 0.9 as in [9]. In *StarGen*-model, we generated the probabilities  $P_L^E$  using the degree distribution of a random preferential attachment graph on the set of  $\ell$  nodes, with incoming nodes attaching with one edge.

Based on our experiments, we chose local  $\alpha$ -values in *StarGen* algorithm at random, uniformly distributed over the interval [1.1, 1.8]. Varying the types of distributions and the boundaries of the sampled interval, we observed that wider intervals

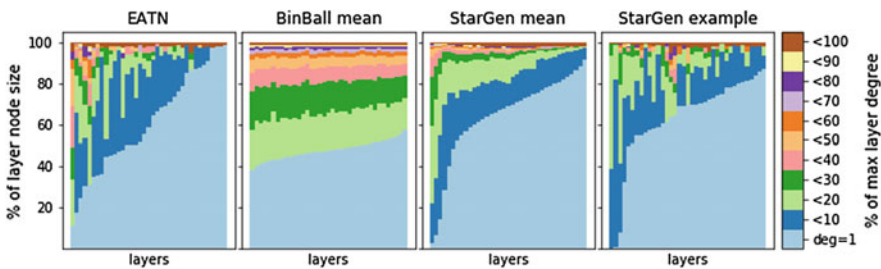
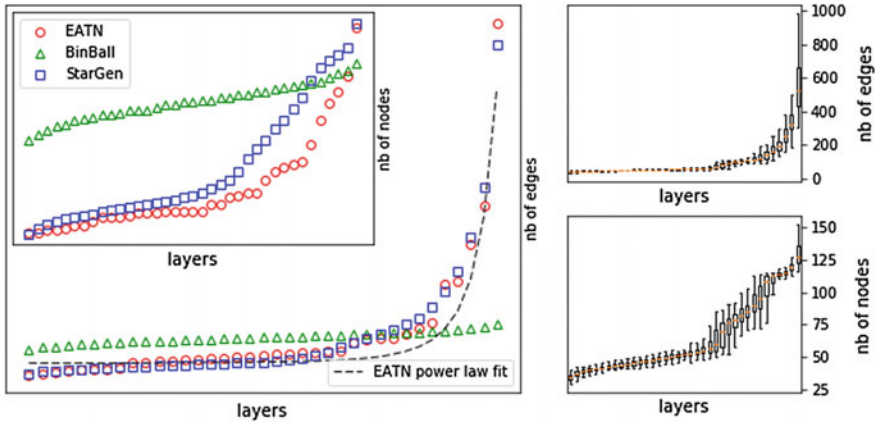


Fig. 1 The comparison of the layer degree structure of the multiplex models

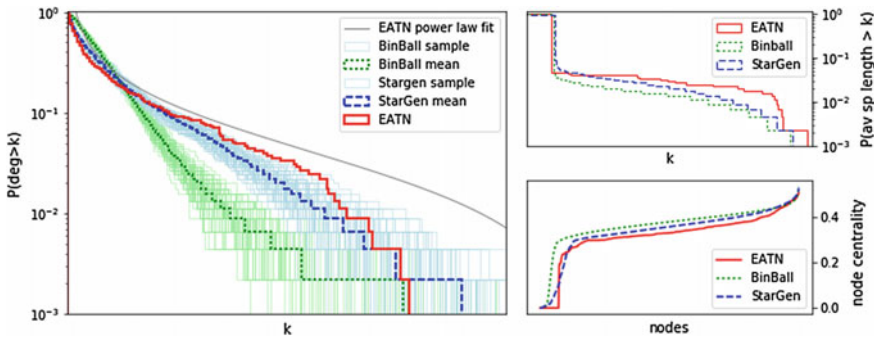


**Fig. 2** Layer edge and node counts comparison: Average over *BinBall* and *StarGen* samples (left), statistics on *StarGen* sample (right)

lead to higher fluctuations of one-degree node count per layer, independently of the distribution. Additionally, the percentage of one-degree nodes increases with growing local  $\alpha$ -values. Therefore we assign small local  $\alpha$ -values to layers with big  $P_L^E$ -values. Furthermore, we noticed that the zero appeal ( $s$ -value) influences the number of zero degree nodes as well as the maximum degree value in the multiplex. In our setting the value  $s = 1.1$  ascertained to perform best.

We refer once more to Fig. 1 showing four plots, the first being EATN, the next one is the average of 100 runs of *BinBall*, followed by the average of 100 runs of *StarGen*, and lastly one example of the analysis of a *StarGen* network. Particularly, the one-degree node count is very large overall and variable for different layers in EATN which we reproduced in *StarGen* due to the varying local  $\alpha$ -value. The other color bands are also less uniform in the *StarGen* than in the *BinBall* samples, and match better the EATN's profile.

Figure 2 shows the edge and node (inset) count per layer for EATN, and the average of 100 runs of *BinBall* and *StarGen* algorithm. The right two figures show the boxplots of the *StarGen* samples. The appropriate choice of the distribution for layer edge counts in *StarGen*-model substantiates the good match of the layer sizes. Even the node sizes evolve adequately, although influenced only by the preferential attachment method and the limit on the maximum value. As seen in Fig. 3, the *StarGen*-model delivers a better model for the EATN-multiplex, based on the degree distribution, the average shortest path length per node, and the average centrality per node. Nevertheless, *StarGen*'s multiplexes tend to come out with higher values for the highest degree nodes.



**Fig. 3** Multiplex: Degree distribution (left), average shortest path length per node (upper right), average centrality per node (lower right)

## 4 Conclusion

Synthetic networks provide a valuable tool to generate replicas of real world networks or to predict their growth. To obtain reliable models, various characteristics of the modeled network have to be reproduced. The more complex the network is, the more challenging it is to design a straightforward procedure to emulate the network. In this work we shaped an easy-to-follow method to replicate a multiplex supporting the variety in the layers' structure. We were able to show that our model considerably outperforms its prototype *BinBall* and delivers a reliable replication of EATN, especially its intra-layer formation.

In our tests we set the interlayer structure out of scope. We observed however that it needs a further consideration as *StarGen*'s as well as *BinBall*'s layers overlap very poorly in comparison with those of EATN.

## References

1. De Domenico, M., Solé-Ribalta, A., Cozzo, E., et al. (2013). Mathematical formulation of multilayer networks. *Physical Review*, *X3*(4). <https://doi.org/10.1103/PhysRevX.3.041022>.
2. Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203–271. <https://doi.org/10.1093/comnet/cnu016>.
3. Boccaletti, S., Bianconi, G., Criado, R., et al. (2014). The structure and dynamics of multilayer networks. *Physical Reports*, *544*(1), 1–122. <https://doi.org/10.1016/j.physrep.2014.07.001>.
4. Barthelemy, M. (2011). Spatial networks. *499*, 1–101. <https://doi.org/10.1016/j.physrep.2010.11.002>.
5. Kim, J. Y., & Goh, K.-I. (2013). Coevolution and correlated multiplexity in multiplex networks. *Physics Review Letters*, *111*(5). <https://doi.org/10.1103/PhysRevLett.111.058702>.
6. Nicosia, V., Bianconi, G., Latora, V., & Barthelemy, M. (2013). Growing multiplex networks. *Physical Review Letters*, *111*(5). <https://doi.org/10.1103/PhysRevLett.111.058701>.

7. Cardillo, A., Gmez-Gardees, J., Zanin, M., Romance, M., et al. (2013). Emergence of network features from multiplexity. *Scientific Reports*, 3(2). <https://doi.org/10.1038/srep01344>.
8. Guimer, R., & Amaral, L. A. N. (2004). Modeling the world-wide airport network. *The European Physical Journal*, 38(2), 381–385. <https://doi.org/10.1140/epjb/e2004-00131-0>.
9. Basu, P., Sundaram, R., & Dippel, M. (2015). Multiplex networks: a generative model and algorithmic complexity. In *IEEE/ACM* (pp. 25–28). <https://doi.org/10.1145/2808797.2808900>.