Network Analysis of MWC 2025 Startup Ecosystem – Josep Cubedo

Executive Summary

This report presents a comprehensive network analysis of startups participating in Mobile World Congress (MWC) 2025. Using natural language processing techniques and graph theory, we analyzed semantic relationships between startups based on their descriptions, revealing distinct industry clusters and potential collaboration opportunities. The visualization highlights how startups in different sectors relate to one another, with particular emphasis on identifying emerging technology trends and cross-industry innovation.

1. Data Collection Process

Methodology

- Source: Direct interviews with company representatives at MWC 2025
- Sample size: 37 startups across multiple technology sectors
- Data points collected: Company name, industry category, and detailed business description
- Validation: Cross-referencing with official MWC directory and company websites

Challenges Encountered

- Description length variation: Some companies provided extensive descriptions while others offered minimal information
- Industry categorization: Many startups operate across multiple sectors, requiring careful primary category assignment
- Language inconsistency: Technical terminology varied widely across companies, necessitating semantic rather than lexical analysis

Data Preprocessing

- Text cleaning: Removed special characters, standardized formatting
- Semantic embedding: Used BERT transformer model to convert text descriptions into 768-dimensional vector embeddings
- Normalization: Applied min-max normalization to similarity scores for consistent comparison

2. Network Construction

Node Definition

- Each node represents a unique startup
- Nodes are colored according to primary industry category
- Node size is consistent to ensure visual clarity

Edge Definition

- Edges represent semantic similarity between startup descriptions
- Threshold value of 0.77 (on a 0-1 scale) used to filter for meaningful connections
- Edge weight (thickness) corresponds to similarity strength
- Curved edges help distinguish multiple connections

Layout Algorithm

- Custom force-directed layout positions similar startups closer together
- Anti-overlap mechanism ensures all nodes remain visible
- Algorithm weights connections based on semantic similarity, creating natural clusters

3. Network Analysis

Topological Characteristics

Key Metrics and Interpretation

The dataset includes **36 startups**, each represented as a node in the network. There are **30 similarity-based connections** (edges) between startups, determined by the similarity threshold of **0.77**.

• **Similarity Threshold**: 0.77. Only startup pairs with a **cosine similarity of 0.77 or higher** are considered connected, ensuring that only **strong semantic relationships** are included.

Network Properties

• **Clustering Coefficient: 0.24**

This indicates a **moderate level of local clustering**, meaning that startups tend to form small groups but not highly interconnected clusters.A **higher clustering coefficient** would suggest that similar startups form tightly knit communities.

• **Average Path Length: 1.63**

The average number of steps needed to travel between connected startups is **1.63**.This suggests that most related startups are **closely linked**, but the network is **not highly interconnected**.

• **Modularity: -0.03**

A slightly negative modularity score indicates that **there is no strong community structure** in the network. This suggests that **similar startups are not forming distinct, well-separated groups** but are instead more evenly distributed.

• **Network Density: 0.0476**

A **low density** means that the **network is sparsely connected**, with only **4.76% of all possible connections** existing.This suggests that while some startups are closely related, **many do not share strong semantic links** based on their descriptions.

• **Network Diameter: N/A (Graph not connected)**

Since the graph is **not fully connected**, there is **no single largest distance** between two startups.This indicates that **some startups have no direct or indirect similarity connections** within the dataset.

• **Average Degree: 1.67**

On average, each startup has **1.67 connections**, meaning that most startups are **only linked to one or two other companies**.This reinforces the idea that **the network is loosely structured** and not highly interconnected.

Key Industry Clusters

1. AI & Machine Learning: Forms the largest and most densely connected cluster
2. Technology & SaaS: Bridges between AI and business-focused startups
3. Business Platforms: Shows significant connectivity with technology clusters
4. Industrial & Robotics: Forms a more isolated cluster with specialized terminology

Bridge Companies

Several startups function as "bridges" between different industry clusters:

- Companies combining AI with business applications show high betweenness centrality
- Startups with diverse technology applications connect otherwise disparate sectors

4. Implications & Conclusions

Strategic Insights

- Partnership opportunities: Highly connected startups within different clusters represent potential strategic alliance targets
- Innovation flow: Bridge companies facilitate knowledge transfer between specialized domains
- Investment focus: Centrally positioned startups may represent emerging technology integration trends

Industry Trends

- AI integration: Widespread embedding of AI across all sectors indicates continued transformative impact
- Sector convergence: Blurring boundaries between traditional technology categories
- Specialized isolation: Highly technical domains like quantum computing remain somewhat isolated

The analysis demonstrates both the diversity and interconnectedness of the MWC startup ecosystem, highlighting how semantic analysis can reveal meaningful patterns beyond traditional industry categorization. The network visualization provides a valuable map for navigating potential partnerships and understanding emerging technology convergence. Different parameters of the functions can be applied to discover new patterns and findings.