Mapping Cancer Metastasis Pathways Using Probability-Weighted Depth First Search on Human Organ Graph

Gabriella Botimada Lubis – 13524006[1,2]
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung
[1]13524006@std.stei.itb.ac.id, [2]lubis.gabriella47@gmail.com

***Abstract*—Metastasis, the spread of cancer cells from a primary tumor to secondary organs, accounts for approximately 90% of cancer-related deaths and remains difficult to map systematically using existing computational approaches. This study proposes representing the human body's organs as nodes in a directed weighted graph, where edges represent clinically known metastasis pathways weighted by probability values derived from the MetMap 500 dataset. A modified Depth-First Search (DFS) algorithm, incorporating probability propagation and threshold-based pruning, is implemented to enumerate and rank all possible metastasis pathways from a given primary organ. Edge weights are computed from the penetrance values reported in MetMap, averaged across all cell lines within each cancer lineage. Experiments conducted on sixteen primary organs show that the choice of pruning threshold creates a clear trade-off between path completeness and clinical relevance, and that pathway rankings vary meaningfully across cancer types, partially aligning with known clinical organotropism patterns. This approach offers a simpler and more interpretable alternative to existing methods such as Bayesian networks or graph homomorphism, requiring only publicly available population-level epidemiological data.**

Keywords—metastasis, mapping, Depth-First Search, directed weighted graph, MetMap, organotropism

I. Introduction

Cancer remains one of the leading causes of death worldwide. According to WHO, the global burden reached 19.3 million new cases and nearly 10 million deaths in 2020. Among the many life-threatening aspects of cancer, the spread of cancer cells from a primary site to secondary organs, known as metastasis, is the most critical factor. Approximately 90% of cancer deaths are caused by metastasis rather than the primary tumor itself[2]. This is supported by statistics showing that cancers detected before metastasis can generally still be treated with survival rates above 90%, which then drop sharply to around 40% or lower once metastasis occurs [6].

One of the greatest challenges in managing metastatic cancer is the limited ability to predict its spreading pathways. Clinically, metastatic pathways are known to be organ-specific, meaning that certain cancer types tend to spread to particular organs in statistically observable patterns[1]. However, systematic mapping of these pathways has generally relied on epidemiological statistical analysis or complex mathematical models such as blood-flow-based differential equations, Bayesian networks, or graph homomorphism.

This paper proposes a simpler and more interpretable alternative, which is representing human organs as nodes in a directed weighted graph, where edges represent clinically known metastasis pathways weighted by spread probabilities derived from historical epidemiological data. On this graph, a modified Depth-First Search algorithm is run to traverse and enumerate all possible metastasis pathways from a primary organ, using a probability-threshold pruning mechanism to filter out clinically insignificant pathways. The resulting pathways are then ranked by cumulative probability to produce a ranked metastasis pathway map.

This paper contributes in two ways. First, it represents the human organ metastasis network as a directed weighted graph constructed from published epidemiological data. Second, it implements weighted DFS with pruning as a traversal mechanism to enumerate and rank metastasis pathways. Unlike Bayesian-based or graph homomorphism methods that require complex cell-sequencing or phylogenetic data, the approach in this paper only requires publicly available population-level metastasis frequency data, making it easier to replicate and interpret.

II. Literature Review

A. Metastasis

Metastasis is the biological process by which cancer cells detach from a primary tumor, enter the circulatory system, and form secondary tumors in distant organs. It is the most dangerous stage of cancer progression and the leading cause of death among cancer patients. This stage occurs through a sequence of steps known as the invasion-metastasis cascade, that are local invasion, intravasation into blood or lymphatic vessels, circulation, extravasation into target tissue, and colonization[2].

Metastasis spreads through two main routes, hematogenous route and lymphatic route. Hematogenous route is where cancer cells travel through the bloodstream to distant organs such as the lungs, liver, brain, and bone. While lymphatic route makes cells spread through regional lymph nodes before reaching secondary

organs. These routes often act in combination rather than independently.

A central phenomenon in metastasis biology is organotropism, the tendency of specific cancer types to spread preferentially to particular organs. This phenomenon first formulated by Stephen Paget in 1889 through the "seed and soil" hypothesis, whish states that cancer cells (the seed) can only thrive in organs (the soil) with a compatible microenvironment. Organotropism is the biological basis for the probability weights used in the graph constructed in this study.

B. Graph Representation for Biological Networks

Graph theory has long been used in computational biology for its ability to formally represent entities and their relationships. A graph $G = (V, E)$ consists of a set of nodes V representing entities and a set of edges E representing relationships between them. In biological contexts, graphs have been used to represent protein-protein interaction networks, metabolic pathways, gene regulatory networks, and disease transmission networks.

For metastasis modeling, the most natural representation is a directed weighted graph, where edge direction reflects the direction of cancer spread and edge weight reflects spread probability. This representation offers some advantages. It intuitively reflects the directional nature of metastasis, since cancer spreads from one organ to another and not symmetrically. It also allows pathway weighting based on measurable epidemiological data, and it enables the direct application of well-established graph traversal algorithms to answer meaningful biological questions, such as determining which spreading pathway is most likely from a given primary organ.

C. Computational Approaches in Metastasis Modeling

Various computational approaches have been developed to model and predict the metastatic process, reviewed here to properly position this study's contribution.

Singh[2] developed a network-based mathematical model representing organs as nodes and blood vessels as edges, simulating cancer cell transport through an advection-diffusion differential equation. This model showed good correlation with clinical data, particularly for gastrointestinal cancers, but requires detailed biological parameters such as blood flow velocity and diffusion coefficients that are not always publicly available.

Margarit et al.[4] used hypergraphs and Markov transition matrices to analyze the distribution of cancer stem cell markers across 23 organs. This study offers a dynamic perspective on marker spread, though focused on marker-level relationships rather than direct organ-to-organ pathways.

Xu et al.[5] performed probabilistic mapping of lymph node metastasis in epithelial ovarian cancer using Bayesian network analysis on retrospective data from 237 patients, producing a detailed probabilistic pathway map to support individualized surgical management, though requiring large, cancer-specific clinical datasets.

Skums et al.[3] applied graph homomorphism to reconstruct metastasis migration history and viral transmission in outbreak settings, which is a powerful but computationally expensive approach for reconstructing evolutionary spread histories.

Staklinski et al.[6] introduced BEAM (Bayesian Evolutionary Analysis of Metastasis), which is a BEAST 2-based probabilistic inference framework supporting full posterior inference over cell lineage trees and migration graphs. This represents the most sophisticated method reviewed, but it depends heavily on cell-lineage tracing data only available from specialized laboratory experiments.

This review reveals a clear gap, as no existing approach explicitly uses graph traversal algorithms to enumerate and rank metastasis pathways based on publicly available population-level epidemiological data. This study fills that gap by proposing weighted DFS with probability-based pruning as a traversal mechanism on a human organ metastasis graph.

D. MetMap as an Empirical Data Source

The metastasis probability data used in this study is drawn from MetMap, a large-scale metastasis mapping resource introduced by Jin et al.[1] and published in Nature in 2020. MetMap was built using an innovative in vivo barcoding strategy, in which a pool of uniquely barcoded human cancer cell lines is simultaneously injected into mouse xenograft models, enabling parallel measurement of metastatic potential across hundreds of cell lines in a single experiment.

MetMap was applied to 500 cancer cell lines spanning 21 solid tumor types, producing a comprehensive organ-specific metastasis map. For each cancer type and target organ combination, MetMap provides a metastatic potential score reflecting the likelihood of that cell line metastasizing to the corresponding organ, covering the lung, liver, brain, bone, and kidney as the primary metastasis sites most frequently studied clinically.

MetMap offers some advantages as a data source for this study, including that its data is generated from controlled and biologically validated experiments rather than mere statistical estimation. Furthermore, its broad coverage across 21 tumor types allows generalization across primary cancer types, and its public availability satisfies the principle of research reproducibility.

E. Depth-First Search Algorithm

Depth-First Search (DFS) is one of the most fundamental graph traversal algorithms in computer science, exploring as far as possible along each branch before backtracking. Starting from an initial node, DFS marks it as visited and recursively visits all unvisited neighboring nodes until no further nodes are reachable, then backtracks to explore unexplored branches. Its time complexity is $O(V + E)$ and space complexity is $O(V)$, making it efficient for moderately sized graphs such as the human organ graph used in this study.

DFS algortihm was selected as the traversal algorithm for several reasons rooted in the biological nature of the metastasis problem. First, the goal of this study is to enumerate all possible spreading pathways from a primary organ rather than finding a

single shortest path to one target, and DFS naturally supports complete pathway enumeration through its recursive backtracking mechanism. Second, metastasis does not spread evenly by distance layer as a Breadth-First Search would model, but instead follows specific and deeper biological pathways. Third, unlike algorithms like A* or Uniform Cost Search, which are optimized to find one optimal path to a single destination, DFS can be easily modified to collect all pathways along with their cumulative probabilities simultaneously. The specific modifications applied to DFS in this study, which include probability propagation and threshold-based pruning.

III. Implementation

A. MetMap Data Extraction

The metastasis probability data in this study is extracted directly from MetMap 500[1], a publicly available dataset generated from an in vivo barcoding experiment on 500 human cancer cell lines. The source file contains six sheets, five of which correspond to the target organs, specifically the brain, lung, liver, bone, and kidney. Each row represents one cell line and contains four measured values, which are the lower and upper confidence interval bounds, the mean metastatic potential score on a logarithmic scale, and the penetrance, which represents the proportion of mice in the experimental cohort that developed metastasis to that organ.

Of these four columns, penetrance was selected as the source of edge weights because, by its mathematical definition, penetrance is computed as the proportion of subjects exhibiting metastasis relative to the total number of subjects in the cohort. This definition is identical to a frequentist probability, making its interpretation as a probability both natural and conceptually direct. This contrast is clear when compared to the mean and confidence interval columns, which represent metastatic potential scores on a logarithmic scale derived from a particular statistical model of the data, measuring the relative intensity or strength of the metastasis signal between cell lines rather than an event proportion.

```
for organ in TARGET_ORGANS:
    ws = wb[f"metp500.{organ}"]
    for row in ws.iter_rows(min_row=2,
    values_only=True):
        cell_line_full, _, _, _, penetrance =
        row[:5]
        lineage = cell_line_full.split("_",
        1)[1]
       data[lineage][organ].append(penetrance)
```

Each cell line in the dataset follows the naming format NAME_LINEAGE, where lineage indicates the primary cancer type of origin, such as A549_LUNG denoting that cell line A549 originates from the lung cancer lineage. All cell lines are grouped by lineage for each target organ.

Of the 22 lineages available in MetMap, this study selects 16 lineages relevant as primary cancer organs, consisting of bone, breast, colon, central nervous system, endometrium, kidney, liver, lung, melanoma, esophagus, ovary, pancreas, prostate, stomach, thyroid, and urinary tract. The skin label is renamed to melanoma since the majority of MetMap skin cell lines originate from melanoma, and the British spelling of oesophagus used by MetMap is standardized to esophagus.

B. Edge Weight Determination

The edge weight between a primary organ and a target organ is computed as the average penetrance value across all cell lines belonging to that lineage. Formally, for lineage l and target organ t:

$$w(l,t) = \frac{1}{n}\sum_{i=1}^{n} penetrance_i$$

where n is the number of cell lines in lineage l tested against organ t. This calculation effectively normalizes the varying experimental responses observed in multiple cell cultures, transforming raw individual data points into a singular representative frequency that serves as a baseline probability for population-level modeling.

```
for lineage, organ_name in
LINEAGE_TO_PRIMARY_ORGAN.items():
   for target_organ in TARGET_ORGANS:
        vals = data[lineage][target_organ]
        if organ_name == target_organ:
            continue
        avg_penetrance = statistics.mean(vals)
        rows.append((organ_name, target_organ,
        round(avg_penetrance, 4)))
```

Edges where the source and target organ are identical, representing self-loops such as liver→liver, are systematically removed from the initial array because an organ cannot biologically be said to metastasize to itself. This rigorous filtering process ultimately yields 76 distinct directed edges originating from 16 primary organs and terminating at MetMap's 5 principal target organs. The resulting dataset is stored directly as a structured CSV file containing three clean columns for the source, target, and probability, which subsequently serves as the core input to graph construction in the next subsection.

C. Graph Construction

The metastasis graph is formally defined as a directed weighted graph:

$$G_m = (V, E, W)$$

where V is the set of relevant human organs, $E \subseteq V \times V$ is the set of clinically known metastasis pathways, and $W : E \to [0,1]$ is a weight function mapping each pathway to its metastasis probability. In this architectural layout, each node represents an organ that can act either as a primary site of tumorigenesis or as a secondary site for colonization, while the directed edges define the chronological and spatial progression of cancer cells from a source organ to a target organ. To programmatically instantiate this network representation, the system reads the formatted CSV file and utilizes the NetworkX library to map the parsed rows into an active directed adjacency structure.

```
df = pd.read_csv(csv_path, comment="#")
G = nx.DiGraph()
for _, row in df.iterrows():
    G.add_edge(row["source"], row["target"],
    weight=float(row["probability"]))
```

An embedded data validation routine is performed immediately during graph instantiation to ensure that every single extracted probability value falls strictly within the mathematically acceptable range between zero and one. This defensive verification step is critical because it guarantees consistency with its interpretation as a true probabilistic weight, which is heavily relied upon by the cumulative probability multiplication inside the DFS algorithm described next.

D. Modified DFS Algorithm

The core algorithm used in this study is Depth-First Search (DFS) modified with two additional mechanisms, which are probability propagation and threshold-based pruning. DFS begins from a single primary organ and recursively traverses all neighboring organs reachable through outgoing edges, known as successors. At each traversal step, the cumulative pathway probability is computed by multiplying the newly traversed edge weight with the cumulative probability of the path so far. Mathematically, for a pathway consisting of k edges denoted as $e_1, e_2, \ldots, e_k$, the joint probability calculation is formulated as:

$$P_{path} = \prod_{i=1}^{k} w(e_i)$$

This dynamic formula updates the running path value continuously as the algorithm steps deeper into the physiological network.

```
for neighbor in G.successors(current):
   edge_weight = G[current][neighbor]["weight"]
   new_prob = cum_prob * edge_weight
```

The threshold-based pruning mechanism halts traversal of a branch completely once its cumulative probability drops below a user-defined numeric threshold denoted as θ.

```
if new_prob < theta:
    continue
```

Since every individual edge weight in this anatomical network is strictly less than one, the cumulative probability always decreases monotonically with each additional step. This specific mathematical behavior means that the pruning mechanism automatically guarantees that the algorithm terminates within a finite number of operations without requiring an explicit depth limit or manual counter.

To prevent traversal from becoming trapped in an infinite cycle, since the organ graph permits recursive feedback loops such as cancer spreading from the liver to the lung and back to the liver, every organ visited along the current active path is tracked using a visited set that is local to that path rather than global across the entire search. This local tracking design allows the same organ to still appear on a different, independent pathway elsewhere in the search tree.

```
def dfs(current, visited, path, cum_prob,
depth):
   if len(path) > 1:
       found_paths.append({"path": list(path),
       "probability": cum_prob})
   for neighbor in G.successors(current):
       if neighbor in visited:
           continue
       visited.add(neighbor)
       path.append(neighbor)
       dfs(neighbor, visited, path, new_prob,
       depth + 1)
       path.pop()
       visited.remove(neighbor)
```

Whenever the algorithm reaches a new organ such that the current path length exceeds one edge, that complete pathway together with its current cumulative probability is stored directly into the global result list. This final output compilation gives researchers a complete and properly sorted overview of every metastasis pathway whose cumulative probability remains above the specified threshold θ.

E. Result Visualization

Once the DFS traversal completes, every discovered pathway is sorted by cumulative probability from highest to lowest to produce a ranked pathway table.

```
df = df.sort_values("probability",
ascending=False)
```

In addition to ranking, the frequency with which each organ appears across all discovered pathways is computed, in order to identify which organs most frequently serve as metastasis destinations from the selected primary organ.

```
for p in paths:
    for organ in p["path"][1:]:
        freq[organ] = freq.get(organ, 0) + 1
```

These results are visualized in three forms, which are the complete organ graph, the same graph with every pathway that passed the threshold θ highlighted in a distinct color per pathway, and a bar chart of organ appearance frequency. All three visualizations are arranged using a circular graph layout in which nodes are ordered alphabetically, ensuring that each organ's position remains consistent across different experimental runs and therefore directly comparable between figures.

IV. RESULTS

A. Graph and Pathway Visualization Results

The primary experiment was conducted using breast as the primary organ with a pruning threshold of $\theta = 0.01$. The results are shown in the figures below.

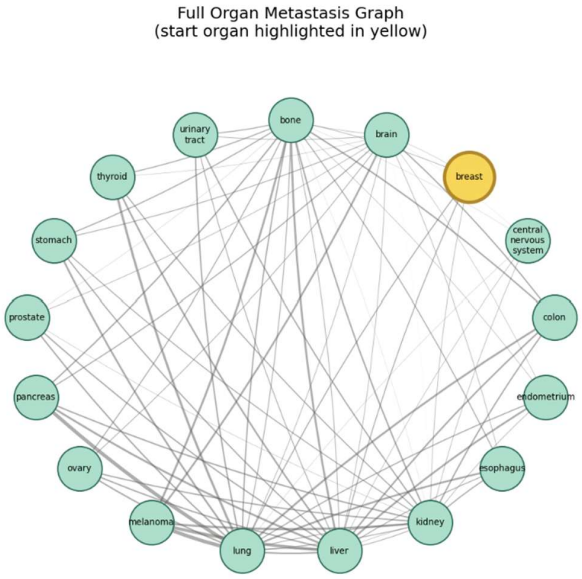

Figure 1. Complete metastasis mapping graph

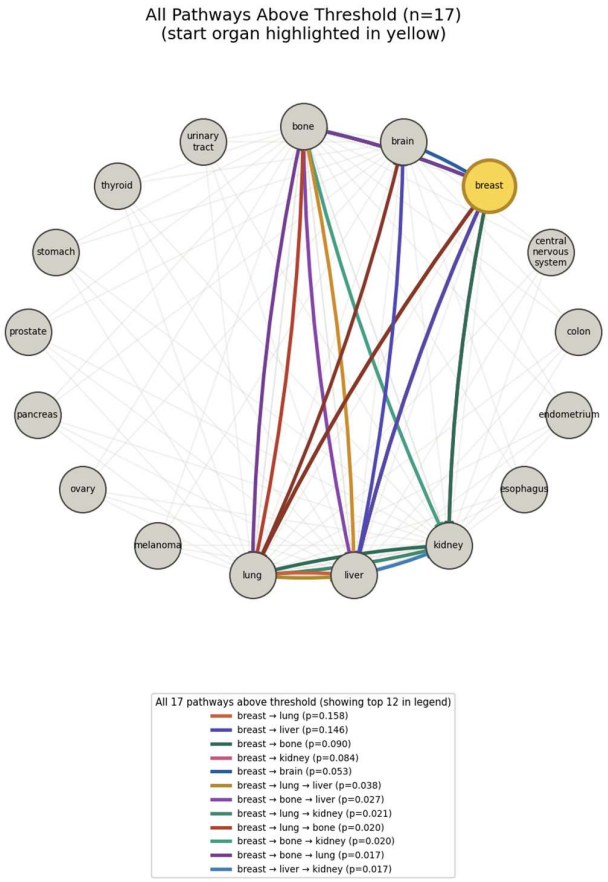

Figure 2. Discovered metastasis pathways above threshold

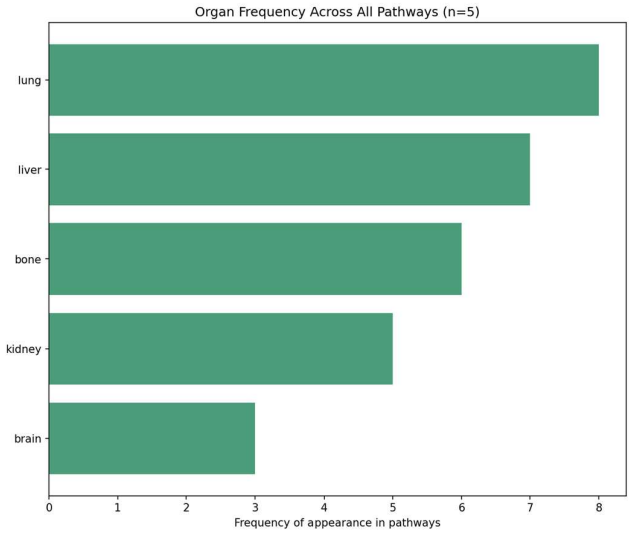

Figure 3. Target organ appearance frequency

Figure 1 shows the complete metastasis mapping graph with the primary organ highlighted in yellow. With this parameter setting, the DFS algorithm identified 17 metastasis pathways whose cumulative probabilities remained above the threshold, as shown in Figure 2. Direct, single-step pathways such as breast to lung, with a probability of 0.158, and breast to liver, with a probability of 0.146, exhibit substantially higher probabilities than two-step pathways such as breast to lung to liver, with a probability of 0.0377. This pattern is consistent with the multiplicative nature of probability in the modified DFS algorithm, where longer pathways necessarily carry smaller cumulative probabilities. Figure 3 shows that the lung is the organ that appears most frequently across all discovered pathways, indicating its role as the primary transit point within the metastasis network originating from breast cancer in this model.

B. Effect of Threshold θ on Pathway Quantity and Characteristics

To evaluate the effect of the pruning threshold, DFS was run from the breast primary organ with five different θ values, specifically 0.05, 0.02, 0.01, 0.005, and 0.001. The results are summarized in the table below.

Table I. Effect of Threshold θ on Pathway Metrics

θ	Number of Paths	Probability Average	Maximum Path Length
0.05	5	0.1062	1
0.02	9	0.0708	2
0.01	17	0.0447	2
0.005	20	0.0389	3
0.001	44	0.0190	3

Table I shows a clear trade-off between completeness and clinical relevance. When θ is increased to 0.05, only 5 pathways pass the threshold, and all of them are direct pathways with the highest average probability of 0.1062. This aligns with

expectations because a strict threshold effectively filters for only the most clinically dominant metastasis pathways. Conversely, when θ is decreased to 0.001, the number of pathways jumps to 44 with the average probability dropping significantly to 0.019, including three-step pathways whose real-world likelihood is very small. These results show that $\theta = 0.01$ used as the default in this stud, provides a reasonable balance by being inclusive enough to capture relevant indirect pathways without flooding the results with negligibly probable ones.

C. Comparison Across Primary Organs

To observe how metastasis pathway patterns differ across cancer types, DFS was run with $\theta = 0.01$ from three different primary organs, namely breast, lung, and melanoma. The results for all three inputs are shown in the figures below, and Table II summarizes the three highest-probability pathways for each primary organ.

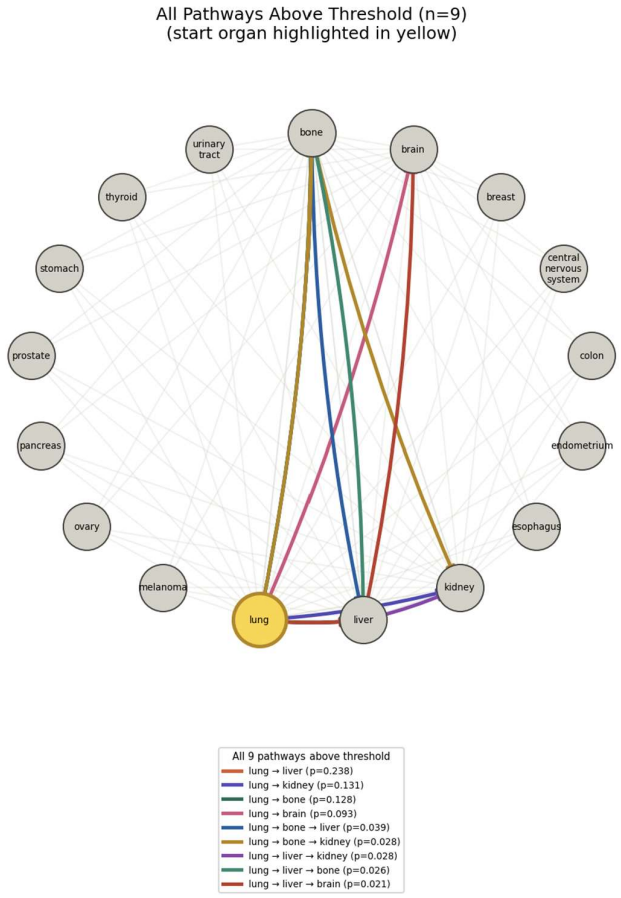

Figure 4. Discovered metastasis pathways above threshold from lung as the primary organ

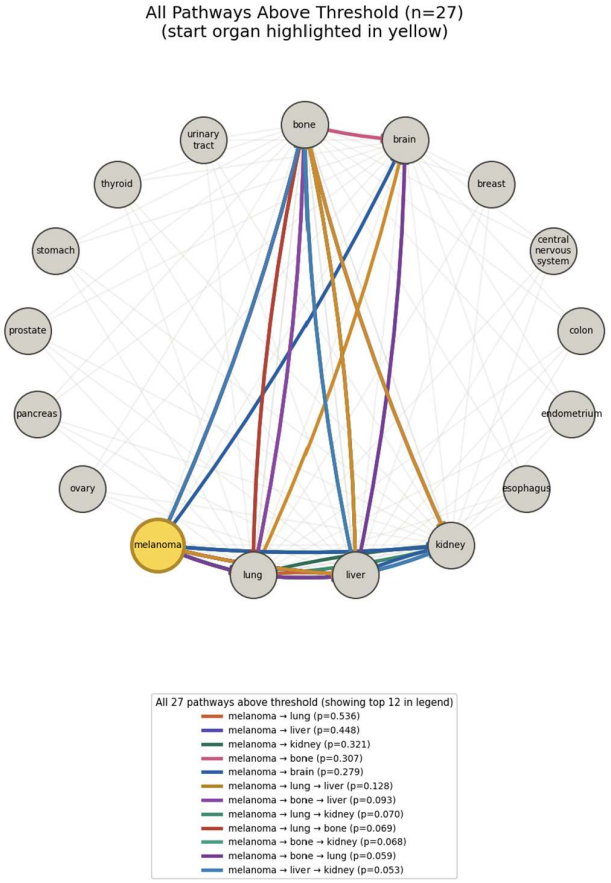

Figure 5. Discovered metastasis pathways above threshold from melanoma as the primary organ

Table II Top 3 Metastasis Pathways Across Different Primary Organs

Primary Organ	Path 1	Path 2	Path 3
breast	breast → lung (0,1584)	breast → liver (0,1460)	breast → bone (0,0901)
lung	lung → liver (0,2383)	lung → kidney (0,1312)	lung → bone (0,1283)
melanoma	melanoma → lung (0,5357)	melanoma → liver (0,4482)	melanoma → kidney (0,3214)

Table II shows that the highest-probability pathway differs across primary organs, where breast most often leads to lung, lung most often leads to liver, and melanoma also most often leads to lung. This difference is consistent with the organotropism phenomenon expected to occur biologically across cancer types, where each primary cancer type has a distinct spreading tendency toward particular target organs. Furthermore, melanoma shows overall much higher metastasis probabilities than breast and lung, as its highest probability reaches 0.5357, which is more than three times the highest probability observed for breast. This aligns well with melanoma's clinical characteristic as one of the most aggressive

cancer types in terms of metastatic potential, spreading widely across multiple organs.

D. Validation Against Known Clinical Patterns

The program's results were compared qualitatively against organotropism patterns extensively documented in clinical literature. Melanoma is clinically known to exhibit broad metastatic tendencies, with the lung and brain as the most frequently affected sites, followed by the liver and bone[7]. In Table II and the complete results in Figure 5, the model produced a specific probability ranking, starting with melanoma → lung ($p = 0.5357$), followed by melanoma → liver ($p = 0.4482$), melanoma → kidney ($p = 0.3214$), melanoma → bone ($p = 0.3071$), and ending with melanoma → brain ($p = 0.2786$).

The ranking produced by the model shows partial agreement with the clinical literature. The lung does occupy the highest position in both sources, which remains consistent between the model and clinical observation. However, the brain, which is clinically described as the second most frequent metastasis site after the lung, instead occupies the lowest position of fifth in this model, while the liver, kidney, and bone rank higher. This discrepancy in brain ranking is likely due to the nature of the MetMap data, which measures colonization potential in animal models through direct bloodstream injection, and therefore does not fully represent the complexity of the blood-brain barrier that clinically affects the rate of brain metastasis in humans. Separately, the breast → bone pathway found in third rank in Figure 2 ($p = 0.0901$) is reasonably consistent with clinical literature that also documents bone as a major metastasis destination for breast cancer, supporting the model's validity for lineages with a larger number of cell line samples, such as breast with 23 cell lines.

E. Limitations of the Results

Based on the program testing conducted, several limitations were identified from the results above.

The first limitation concerns the imbalance in the number of cell lines across lineages in the MetMap dataset used as the source of edge weights. The number of cell lines per lineage ranges from 98 for lung to as few as 4 for prostate, with several other lineages such as thyroid with 10 and bone with 13 also classified as low. This imbalance directly affects the statistical reliability of the resulting edge weights, since an average penetrance computed from only 4 cell lines is far more susceptible to random variation than one computed from 98 cell lines. This effect is already visible in the brain-ranking discrepancy observed in the melanoma results in Subsection D, given that brain is a moderately represented lineage with 35 cell lines compared to lung with 98 cell lines, which instead emerged as the target organ with the highest probability. Discrepancies of this kind are most likely not caused by errors in the DFS algorithm itself, but rather by the limitations and imbalance of sample sizes in the underlying data source.

The second limitation lies in the static nature of the constructed graph, which does not model the time dimension at all. Edge weights represent aggregate metastasis probability across the entire tested cell line population without distinguishing whether spread occurs within weeks, months, or years after the primary tumor emerges. As a result, two metastasis pathways with identical cumulative probability in this model may correspond to very different real-world clinical courses, where one progresses acutely and aggressively while the other develops slowly over years. This progression-speed information is entirely uncaptured by the probabilistic graph-based model used in this study, meaning the results should be understood strictly as a map of pathway likelihood rather than a prediction of event timing.

The third limitation is the restricted coverage of target organs, which is limited to only the five organs experimentally measured by MetMap, specifically the brain, lung, liver, bone, and kidney. This restriction is a direct consequence of MetMap's own experimental scope, which measured metastatic potential to only these five organs across all 500 tested cell lines. Consequently, metastasis pathways to other clinically common destinations, such as the adrenal glands, peritoneum, or skin, cannot be represented by this model at all, even though these organs are highly relevant metastasis sites for certain cancer types.

V. Conclusion

This study demonstrates that human cancer metastasis pathways can be effectively mapped and ranked using a modified Depth-First Search algorithm operating on a directed weighted graph constructed from publicly available MetMap data. By combining probability propagation with threshold-based pruning, the proposed method enumerates clinically meaningful pathways while avoiding the computational and data complexity required by more sophisticated approaches such as Bayesian networks or graph homomorphism. Experiments across multiple primary organs and pruning thresholds reveal a consistent trade-off between pathway completeness and clinical relevance. They also show that the resulting pathway rankings partially align with known organotropism patterns, while simultaneously surfacing meaningful discrepancies that point to specific limitations in the underlying data rather than the algorithm itself. The simplicity and reproducibility of this approach, which requires only public and population-level metastasis frequency data, make it a practical complement to more data-intensive computational metastasis models, offering clear potential value as an exploratory tool in cancer research and biomedical education.

VI. Appendix

Github: https://github.com/gabriellaalubis/Weighted-DFS-on-Cancer-Metastasis-Pathway-Mapping

VII. Acknowledgement

The author thanks God Almighty for His blessings and guidance that allowed for the completion of this paper. The author deeply appreciates the School of Electrical Engineering and Informatics at Institut Teknologi Bandung and Ir. Rila Mandala, M.Sc., Ph.D as the lecturer of the Strategi Algoritma course for the mentorship and guidance. Finally, the author expresses gratitude to her family and friends for their continuous encouragement and motivation throughout the study.

REFERENCES

[1] L. Jin et al., "A metastasis map of human cancer cell lines," Nature, vol. 588, 2020.

[2] T. Singh, "Mathematical modelling on the network of metastasis," Ph.D. dissertation, Univ. of the Witwatersrand, Johannesburg, South Africa, 2022.

[3] P. Skums et al., "Reconstruction of the evolutionary history of metastasis and pathogen transmission via graph homomorphism," Nature Communications, vol. 16, 2025.

[4] M. Margarit et al., "Hypergraph analysis of cancer stem cell marker distribution across human organs," arXiv:2407.19330, 2024.

[5] J. Xu et al., "Bayesian network analysis of lymph node metastasis distribution in epithelial ovarian cancer," PMC, 2026.

[6] S. Staklinski et al., "BEAM: Bayesian Evolutionary Analysis of Metastasis," bioRxiv, 2025.

[7] O. Abdel-Rahman, "Clinical correlates and prognostic value of different metastatic sites in patients with malignant melanoma of the skin: a SEER database analysis," Journal of Dermatological Treatment, vol. 29, 2018.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 19 Juni 2026

Gabriella Botimada Lubis
13524006