

Cápsula 4: Matriz de adyacencia

Hola, bienvenidxs a una cápsula del curso Visualización de Información. En esta hablaré sobre el *idiom* de matriz de adyacencia, que es otra forma de representar visualmente datos de red.

Específicamente, se organiza de forma que todos los nodos de la red se distribuyen en los bordes horizontal y vertical, y las combinaciones de celdas correspondientes representan un posible enlace entre nodos mediante una marca de área. Esta indicaría si existe una conexión usando el color del área como canal.

La codificación más básica permite distinguir si hay conexión o no, mediante presencia o no de color. También es posible y se suele codificar un atributo adicional de los enlaces utilizando un *colormap* de relleno de las marcas. Así es posible mostrar no solo conexión, sino que también alguna magnitud o identidad de la conexión, para redes etiquetadas o con pesos.

Para redes no dirigidas las matrices son simétricas, al presentar conexiones en ambos sentidos. En esos casos se pueden solo mostrar la mitad de la matriz, la que esté sobre la diagonal o bajo ella. En cambio para redes dirigidas es necesario mostrarla completa.

Las matrices de adyacencia logran alta densidad de información, ya que con suficiente espacio podría llegar a los miles de nodos e incluso un millón de enlaces. Esto es común con representaciones matriciales como los mapas de calor matricial que son densos y compactos.

En comparación al *idiom* de nodo-enlace revisado en la cápsula anterior, este último le gana por bastante en popularidad a la matriz de adyacencia como método de representación de redes.

La gracia de enlace-nodo es que para redes suficientemente pequeñas la representación es muy intuitiva para muchas de las tareas abstractas que aparecen con datos de red. Sobre todo se lucen para tareas relacionadas con entender la topología de la red, como identificación de caminos y detección de vecindades alrededor de nodos.

La limitante que aparece en enlace-nodo es su límite de escala al considerar redes de múltiples nodos y de densidad de enlace grande. Se vuelven imposibles de ver por oclusión entre conexiones que cubren otros nodos.

Incluso con densidad de enlaces baja, muchos nodos presentes degeneran rápidamente el espacio de visualización. Por eso hay mucho esfuerzo hoy en día en mejorar el procesamiento y desarrollo de técnicas para visualizar redes grandes.

De forma alternativa, la matriz de adyacencia escala bien tanto con número de nodos y densidad de enlace. En esta no existe oclusión producto de la presencia de enlaces, por lo que puede albergar espacio para redes poco densas como completamente conectadas incluso.

Otros beneficios de las matrices es que son predecibles en estructura, al tener una forma simple y fácilmente reproducible, son fáciles de agregar a la composición de una herramienta, a diferencia de los nodo-enlace, cuyo espacio ocupado puede variar y ser ineficiente en uso.

También son estables, al agregar un elemento nuevo, el cambio visual es menor, mientras que en enlace-nodo puede provocar cambios importantes. Eso les permite ser fáciles de juntar con *idioms* de navegación y son compatibles con el reordenamiento de sus filas y columnas. Esto último mejora la experiencia de búsqueda de nodos específicos, algo bastante difícil para el *idiom* de enlace-nodo básico.

El contra de matrices de adyacencia es que son más difíciles de leer, y menos intuitivas de ocupar para el usuario común. Pero, con entrenamiento suficiente, muchas tareas de redes pueden volverse fáciles. Como detectando patrones de cúmulos de nodos interconectados. Incluso hay diferenciaciones entre patrones, que llevan a notar casos más específicos.

Pero tal vez peor aún, las matrices comunican topología de manera muy indirecta. Aprender la identificación de caminos es mucho más difícil que en enlace-nodo. Han habido estudios que sugieren que el *idiom* de enlace-nodo es mejor para redes pequeñas, mientras que la matriz de adyacencia es mejor para redes grandes en general.

Específicamente, tareas como: estimación de número de nodos y enlaces, encontrar los nodos más conectados, encontrar un nodo por etiqueta y encontrar un enlace entre un par de nodos; se vuelven más difíciles para enlace-nodo a medida que crece la red. En cambio en matrices de adyacencia la eficacia prácticamente no cambia.

Y con eso termina el contenido de esta cápsula. Recuerda que si tienes preguntas, puedes dejarlas en los comentarios del video para responderlas en la sesión en vivo de esta temática. ¡Chao!