Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Social Networks in the Upin & Ipin Universe: Comparative Analysis of a Series and Movie in Southeast Asian Media

Alyssa April Dellow
Universiti Kebangsaan Malaysia
Fatimah Abdul Razak
fatima84@ukm.edu.my
Universiti Kebangsaan Malaysia

Article

Keywords: social network analysis, character interactions, Upin and Ipin, character network, Malaysian animation

Posted Date: June 9th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6329035/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

SOCIAL NETWORKS IN THE UPIN & IPIN UNIVERSE: COMPARATIVE ANALYSIS OF A SERIES AND MOVIE IN SOUTHEAST ASIAN MEDIA

ABSTRACT

Social network analysis offers a unique perspective into media narratives by examining character interactions. This study explores character prominence and narrative flow in the Upin & Ipin series, including Seasons 1 and 2, and the first Upin & Ipin movie, "Geng: The Adventure Begins". Speaker-listener identification is applied to generate directed and weighted character networks. Findings reveal that the character networks mimic real-life social patterns, showcasing communal network structures over more egocentric structures commonly observed in Western media. High betweenness centrality characters serve as superspreaders, bridging central and peripheral groups, while antagonists have limited centrality, suggesting minimisation of hazardous behavioural emulation in young audiences. The movie exhibits more intricate web interactions than the series, with deliberate character placements indicating changes in narrative strategies. These results demonstrate the utility of social network analysis in uncovering hidden character dynamics beyond conventional storytelling analysis, in addition to laying the foundation for media comparative studies in Southeast Asia.

Keywords: social network analysis; character interactions; Upin and Ipin; character network; Malaysian animation

1. Introduction

Television serves as a powerful medium that encapsulates a vast variety of human knowledge via stories embedded in character relations (Nan et al. 2015). An interesting aspect of the Malaysian film industry lies within the role of voice actors as connectors between actors from diverse genres, seeing as their appearances extend across other film genres (Bahari et al. 2023). This highlights the significance of Malaysian animation not only as a niche market, but also as a link between various genres in Malaysia's cinematic landscape. Malaysia's animation industry has gained recognition with locally produced films such as Upin & Ipin, BoBoiBoy, Ejen Ali, Mechamato and Didi & Friends. Among them, Upin & Ipin remains Malaysia's pride, achieving commercial success locally and internationally. The first season aired in 2007 during Ramadan (the ninth month of the Muslim year during which Muslims keep a rigorous fast from dawn till sunset) with six episodes on TV9 (Fandom n.d.-b), initially aiming to educate kids on appreciating the holy month. Now spanning 18 seasons and three movies, Les' Copaque Production has turned Upin & Ipin into a household name across Southeast Asia (SEA), making its way into Indonesia (285.1 million population), Singapore (5.9 million population), Brunei (465k population), India (1.4 billion population) and other countries, even streaming on Disney and Netflix in United Kingdom and Australia (Zeanaaima 2022). Beyond films, the Upin & Ipin franchise continues to evolve into interactive media with a planned open-world game for PlayStation 4 and Nintendo Switch, developed in collaboration with Streamline Studios (Weiringen 2024). This move reflects Les' Copaques ambition for global expansion.

Past research on Upin & Ipin primarily focused on its themes and impact. For instance, Yulianeta and Amandangi (2021) explored gender issues, while Ribka Yulista Devi Purnamasari's (2020) highlighted how kids aged seven to nine imitate the characters' language behaviours. Social network analysis (SNA) leverages graph theory to extract information from films, series, novels or movies. By generating networks from character interactions, the tangible form of relationships may bear insight into the complex web of links that could otherwise be

missed by other approaches (Schauf & Escobar Varela 2018). This study aims to generate character conversational networks to compare Seasons 1 and 2 of the Upin & Ipin series against its first movie, "Geng: The Adventure Begins". These networks reveal key characters, hidden and structural roles, offering an overview of how characters facilitate or disrupt relationships that drive the story's narrative. This relates to "role theory", which holds an individual's positional role accountable for their behaviours (Biddle 1986). Given animation's significant role in developing viewpoints (Fonseca et al. 2023), Upin & Ipin offers an opportunity to introduce Malaysia's multicultural landscape to local and global audiences.

SNA serves as a powerful framework for investigating connections across fields, including entertainment (Giri et al. 2024), marketing (Zaefarian et al. 2022), telecommunications (Mandić et al. 2018) and education (Abdul Razak et al. 2019; Rahim et al. 2023; Ratanarajah et al. 2020). It helps uncover how relationships create perception biases (E. Lee et al. 2019) and enhances media analysis beyond audiovisual metrics (Yeh & Wu 2014). Park et al. (2012) and Jung et al. (2013) found that analysing character networks reveal a multitude of narrative insights, as relationships build the bulk of a story's structure. As characters reflect the intentions of producers in creating a story, they act as the foundation for any form of literature (Chowdhury et al. 2019). Not only do character networks aid in understanding characters and narratives (Murakami et al. 2020), but they also contribute to a media's popularity (Sugishita & Masuda 2023).

Co-occurrence analysis is widely used for detecting character interactions due to its simplicity (Labatut & Bost 2019). It divides media into smaller units and records interactions between character pairs when they appear together within the same unit. Despite their advantages, co-occurrence networks are merely stand-ins for actual interactions that lack directionality, providing limited insights (Labatut & Bost 2019). Characters may share scenes without interacting by silently witnessing an incident (Min & Park 2016), potentially inflating the importance of minor characters (Prado et al. 2016).

Conversational networks, which track explicit verbal interactions, offer a more sophisticated alternative. It is more challenging to implement since it requires determining both speakers and listeners (Labatut & Bost 2019). Anyhow, conversational networks provide deeper insights into structural patterns by enabling the extraction of directed networks (Park et al. 2012). They are a more refined subset of co-occurrence networks, as all conversing characters co-occur, but not vice versa (Bost et al. 2016). Conversational networks work best for dialogue heavy media (J. Lee & Yeung 2012)like Upin & Ipin, which is rich in dialogues but have minimal inner monologues or non-verbal tension. Previous SNA studies have analysed various media formats individually, including:

- Movies: You've Got Mail (Weng et al. 2009), The Devil Wears Prada (Yeh & Wu 2014)
- Series: Stargate and Star Trek (Tan et al. 2014), Friends (Nan et al. 2015), Game of Thrones (Bost et al. 2019)
- Comics: Japanese manga (Murakami et al. 2020; Sugishita & Masuda 2023), Marvel comics (Alberich et al. 2002; Gleiser 2007; Shi et al. 2021)

The Upin & Ipin universe thrives on its diverse cast, each contributing to the story via distinct personalities and communication. Extended Data Fig. 1 (a) and (b) present alternative network versions from Fig. 1, whereas Extended Data Tab. 1 details character appearances, roles, personalities, genders and races. The show revolves around the lives of five-year-old orphaned twin boys, who live with their elder sister, Kak Ros, and grandmother, Opah, in Kampung Durian Runtuh. Grounded in a realistic Malaysian village life, the series emphasises morals, family, religion, cultural identity and community, appealing to young audiences

through relatable characters and lighthearted humour. Seasons 1 and 2 follow the twin's Ramadan experiences, as they learn about fasting, prayer, respect and discipline. Guided by Opah and Kak Ros at home, and Cikgu Yasmin (Teacher Yasmin) at Mesra Kindergarten, they navigate Ramadan with pals like Ehsan, Fizi and Mail, culminating in the joyous Eid al-Fitr celebration, which marks the end of Ramadan.

In 2009, the Upin & Ipin franchise extended to the big screen with "Geng: The Adventure Begins", an action-packed blend of real-world events and fantasy. The movie begins with Badrol and Lim returning to Kampung Durian Runtuh after Tok Dalang's (Badrol's grandfather) interview about a mysterious Durian Ghost responsible behind the village's missing durians. The duo is later captured by the alleged ghost, revealed to be Singh, an antagonist. Meanwhile, Upin, Ipin, Kak Ros and Rajoo chase Oopet, a strange creature secretly eating the missing durians. They are then led to the poachers' (Pak Mail, Singh, Salleh) hideout, where Oopet's mother, along with Badrol and Lim are held captive. The movie concludes with Oopet and his mother returning to their world, while the youngsters reunite with the village elders (Tok Dalang, Opah, Muthu, Ah Tong).

"Geng: The Adventure Begins", Malaysia's first CGI animated feature film, was initially envisioned as the main feature to launch the Upin & Ipin universe (ASEAN Intellectual Property Portal n.d.). The series was originally a side project planned before the movie's launch to gauge audience acceptance to a uniquely Malaysian animation. Upin and Ipin, though not the movie's main focus, became breakout stars by its release, shifting public perception to assume that they were always the main protagonists. Making Upin and Ipin the series stars was a practical choice, since modelling bald twins simplified animation production complexity. What began as a promotional project led to massive success, with Season 2 achieving local viewership of 1.5 million, surpassing even SpongeBob SquarePants (Fandom n.d.-b). Rather than launching a new project, producers turned the two-season series into a long-running TV series. In contrast, the movie's other major characters are yet to receive follow up projects. Nonetheless, the movie performed impressively well at the box office, collecting 6.31 million Ringgit in Malaysia throughout its seven-week screening, besides securing international releases in Indonesia and India (Fandom n.d.-a).

The shift in focus shaped the future trajectory of Malaysian animation. The second movie, "Upin & Ipin: The Lone Gibbon Kris", introduced more characters to expand the Upin & Ipin universe, akin to the Marvel Cinematic universe (MCU) and Detective Comics universe (DCU). These superheroes have become cultural icons in the Western community (Alberich et al. 2002), and arguably on a global scale. The shared universe concept allows character crossovers across different films (Gleiser 2007), which could lead to side characters starring in standalone spin-offs. An example from the animation industry perspective is Puss in Boots, a Shrek spin-off by DreamWorks. Likewise, Les' Copaque may be adopting the same strategy, paving the way for potential spin-offs. A prime example is Ultraman Ribut, introduced in the eighth season of Upin & Ipin in 2014 via a collaboration with Tsuburaya Productions, producers of the iconic Ultraman franchise (Azizul Rahman Ismail 2025). Incorporating silat (a martial art originating from SEA) inspired combat, Ultraman Ribut often appears in Ultra Galaxy Fight: New Generation Heroes, an Ultraman spin-off mini-series, marking a milestone for Malaysian animation. Although the MCU and DCU possess international recognition, Upin & Ipin have emerged as cultural icons in SEA, demonstrating the cross-border value of a well-developed universe.

The widespread appeal for this animation raises key questions: "Can Upin & Ipin aid in understanding SEA storytelling narratives that help shape human behaviour?", "What drives its success, and can it be replicated?", "Can SNA quantify narratives beyond traditional analysis?".

Despite its popularity, systematic analysis of character interactions and narrative structures in SEA animation remains scarce. SNA bridges this gap by extracting insights beneficial to animators, scriptwriters and cultural researchers. By analysing the Upin & Ipin series (Seasons 1 and 2) and "Geng: The Adventure Begins", this study provides a novel perspective on Malaysian animation, contributing to a broader understanding of SEA's animation culture and supporting the development of the country's entertainment industry. Although regionally focused, the findings are globally relevant, demonstrating their applicability across foreign media since no prior research has applied SNA to compare two media formats, specifically a series and a movie from the same animated franchise.

2. Results

2.1. *Character Conversational Network: Structure, Community Detection and Global Statistics*

From Tab. 1, the movie introduces 11 additional characters compared to the series, increasing connections in both directed and undirected networks. This suggests that animated movies have larger, more diverse and potentially more complex networks despite comparable runtimes. The higher global clustering coefficient (GCC) in the series indicates a tightly knit group where characters are more likely to interact within their groups, while the movie characters are more dispersed, as seen in the community detection of Fig. 1. The series has two large communities, with Salleh and Ayah Ehsan (Ehsan's father) being on their own, whereas communities are more distributed in the movie. This aligns with the series' average path length of 1.484, reflecting a more efficient information flow as most characters are directly connected. In the movie, intermediaries are likely needed to bridge two characters as they belong to different clusters, highlighting a more layered and complex storyline. Messages travel within 1 to 2 steps in the series and 2 to 3 steps in the movie on average. According to the Six Degrees of Separation theory that arose from Karinthy's (1929) short story and later explored by Milgram (1967), everyone in a network is only six or fewer steps apart. Hence, the series and movie networks exhibit small average path lengths, a characteristic also observed in Shi et al.'s (2021) study on Marvel comics. In addition, the GCC values in both networks exceed their respective densities, indicating that the GCC is high. The combination of small average path lengths and high GCC values are characteristics of small-world networks, a hallmark of most real-world social networks (Watts & Strogatz 1998).

Tab. 1 | Global network statistics

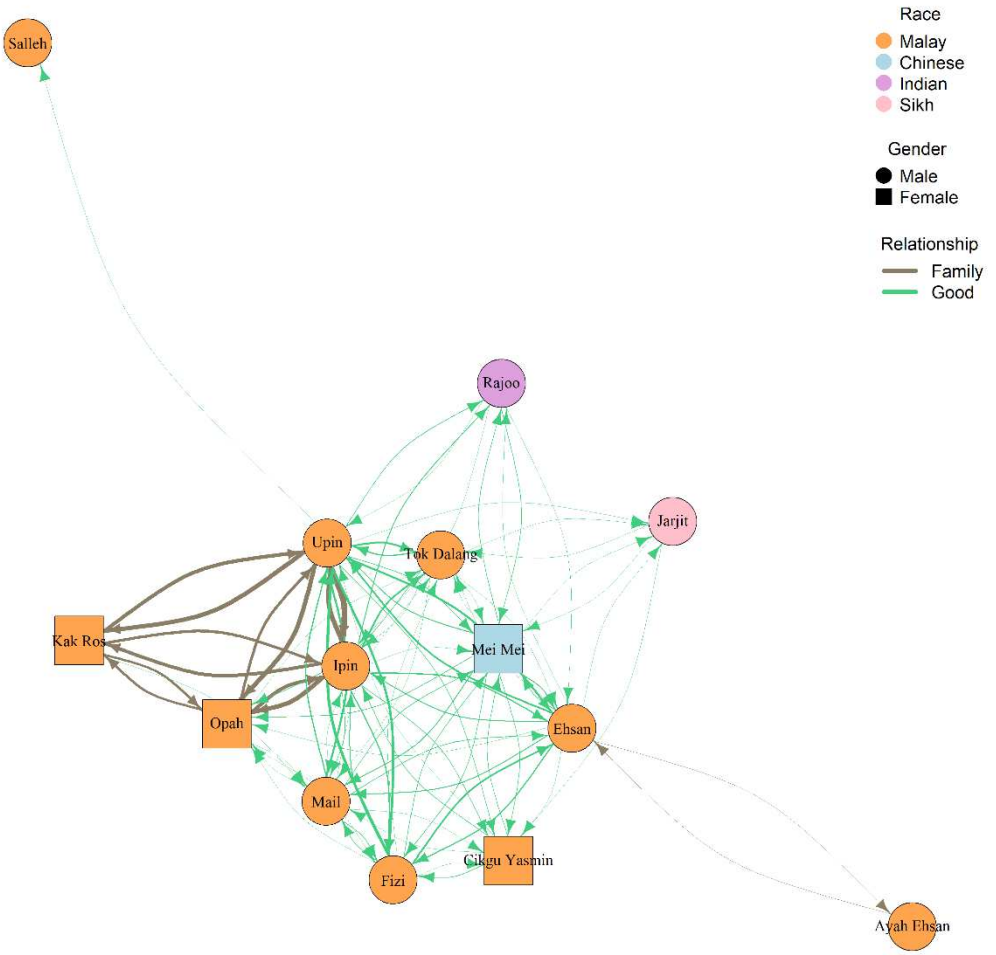
	Series	Movie
Number of Nodes	14	25
Number of Edges (Directed)	88	114
Number of Edges (Undirected)	49	69
Global Clustering Coefficient	0.738	0.525
Average Path Length	1.484	2.126
Density	0.538	0.230

Since network visualisation is an important first step in SNA (Rahim et al. 2023), Fig. 1 (a) and (b) portray the directed and weighted character networks for the series and movie respectively. Each plot conveys multiple layers of information, node colours represent race, node shapes indicate gender, edge weights reflect relationship strength and edge colours distinguish between good, bad or familial ties. Animals are categorised under race with yellow

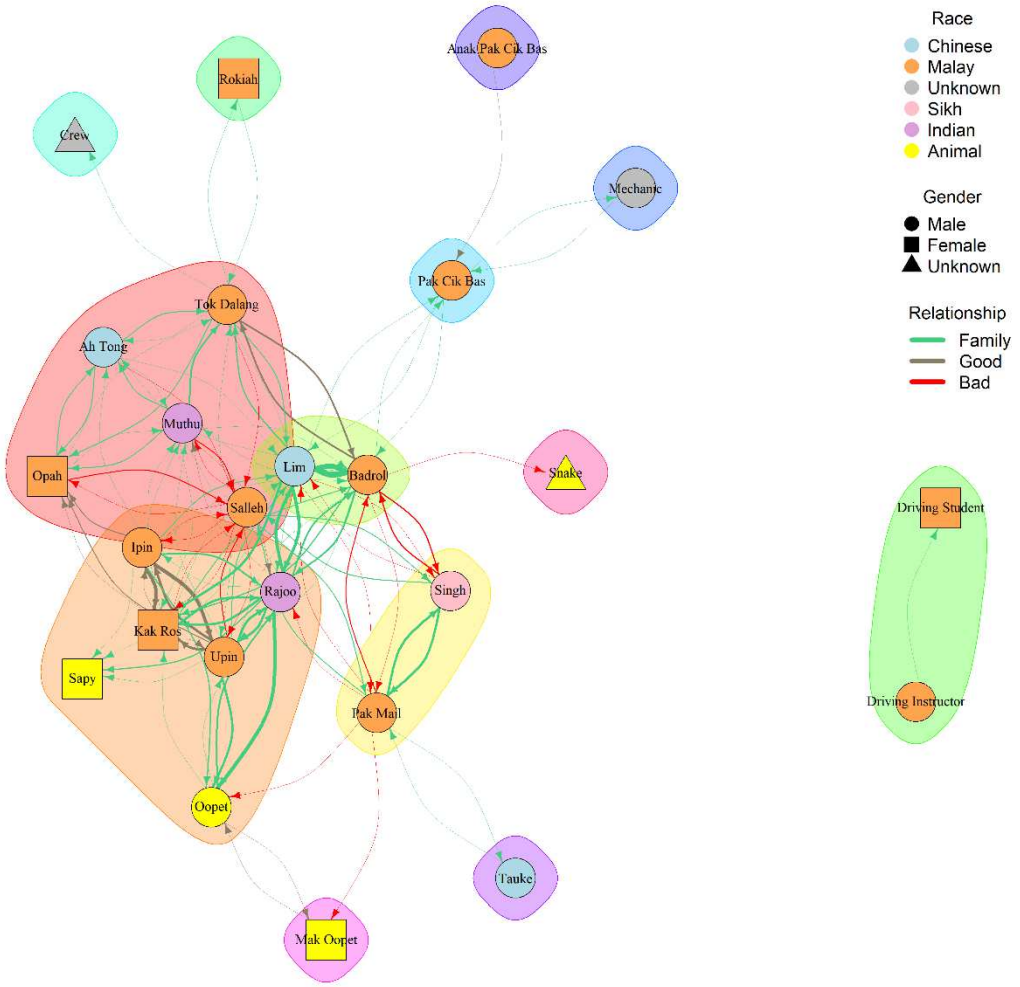
nodes, while unknown gender applies to animals with unspecified genders, or individuals who appear only through off-screen interactions and phone calls. The Walktrap community detection method defines character clusters based on denser intra-cluster and sparser inter-cluster connections.

The series predominantly teaches children valuable lessons about morality, family, religion and friendships, as there is a distinct division between the family (Upin, Ipin, Kak Ros, Opah), and the other cluster consisting of friends, the school teacher and other villagers of Kampung Durian Runtuh. This demonstrates the series' themes, with the family cluster focusing on moral dialogues, while the other cluster emphasises friendships in school and among villagers. Clearly, the first two seasons stress on familial relationships, as the conversations between Upin, Ipin, Kak Ros and Opah are more frequent.

The movie introduces new relationships and expands on existing ones. The largest cluster includes younger characters and animals, likely due to their shared adventure and Rajoo's ability to communicate with animals. The cluster of five including Tok Dalang, Muthu, Ah Tong, Opah and Salleh represent older characters. Despite Salleh's association in poaching with Pak Mail and Singh, he remains in the older character cluster due to his villager status, making his betrayal more pronounced. This shows that Pak Mail and Singh are closer and have distinct roles than Salleh, with the introduction of antagonists adding complexity and excitement to the movie. Furthermore, Badrol and Lim, who are best friends, remain in their own circle due to their frequent conversations. The separation of protagonists from antagonists, as well as younger from older characters, contributes to the narrative as each character plays their part. Notably, both the series and movie reflect good cultural practices as characters of different races and genders mix around as a community.


Race
Malay
Chinese
Indian
Sikh
Gender
Male
Female
Relationship
Family
Good
Salleh
Rajoo
Jarjit
Upin
Tok Dalang
Kak Ros
Mei Mei
Ipin
Opah
Ehsan
Mail
Cikgu Yasmin
Fizi
Ayah Ehsan

(a)



(b)

Fig. 1 | Directed and weighted (a) series and (b) movie network plots with Walktrap community detection

2.2. *Centrality Measures*

Since character interactions vary between the series and movie, examining centrality measures helps comprehend each character's prominence within the network. The top section of Tab. 2 lists five characters with the highest degree centralities (DCs) in the directed and weighted networks, considering both the number of connections and interaction frequency. In case of centrality value ties, more than five characters may be listed.

In the series, Upin and Ipin's status as the main protagonists is reinforced, as they hold the highest DCs at 0.228 and 0.176, followed by Opah and Kak Ros. This portrays that the series primarily revolves around family. In the movie, however, the narrative seems to place a greater focus on Rajoo, Lim and Badrol who exhibit higher interaction frequencies with other characters, surpassing Upin and Ipin's prominence.

A glance at the undirected and unweighted centrality outcomes for the series reveals a consistent set of central characters across different metrics. Upin, Ehsan, Ipin, Mei Mei and Mail appear in the top five for all centrality metrics, underscoring their central roles. Opah ranks high in degree and closeness centralities, portraying her many connections and proximity to others. Interestingly, Kak Ros does not rank in the top five for any undirected and unweighted centrality measures in the series despite her frequent interactions in Fig. 1 (a). Her presence in the top five for directed and weighted DC comes from repeated interactions within her family. Since most of her scenes take place at home, her interactions with the broader community are limited. As a result, this restricts her from being in close contact with others, besides reducing her ability to bridge different characters. Kak Ros primarily acts as a parental figure to her younger brothers.

As for the movie, character prominence shifts to those not from Kampung Durian Runtuh, such as Badrol and Lim, as well as previously absent or small characters in Seasons 1 and 2 of the series like Muthu, Pak Mail and Pak Cik Bas (bus driver) and Rajoo. The series exhibits more nuanced centrality values across all metrics except betweenness, which is much higher in the movie. This accentuates the importance of certain small characters such as Tok Dalang, Pak Mail and even Pak Cik Bas as critical intermediaries that bridge isolated parts of the movie network. Nonetheless, with the highest eigenvector centrality (EC) scores, Ipin and Upin maintain strong ties to the network's central figures. Interestingly, Upin is more central in the series, whereas Ipin takes a bigger role in the movie. Besides that, Singh does not rank highly in any centrality measure, while Pak Mail ranks fourth or fifth despite them being the main antagonists.

Tab. 2 | Top five degree centrality values for the directed and weighted networks, along with the top five degree, eigenvector, betweenness and closeness centrality values for the undirected and unweighted networks. Brackets next to character names indicate gender: (M) for male and (F) for female.

	Directed and Weighted Degree Centrality			
	Series		Movie	
1.	Upin (M)	0.228	Rajoo (M)	0.128
2.	Ipin (M)	0.176	Lim (M)	0.118
3.	Opah (F)	0.116	Badrol (M)	0.117
4.	Kak Ros (F)	0.097	Upin (M)	0.094
5.	Fizi (M) & Ehsan (M)	0.077	Kak Ros (F)	0.083

	Undirected and Unweighted Degree, Eigenvector, Betweenness and Closeness Centrality							
	Degree		Eigenvector		Betweenness		Closeness	
Series								
1.	Upin (M)	0.122	Upin (M)	1.000	Upin (M)	17.826	Upin (M)	0.071
2.	Ehsan (M)	0.112	Ehsan (M)	0.954	Ehsan (M)	14.743	Ehsan (M)	0.067
3.	Ipin (M) & Mei Mei (F)	0.102	Mei Mei (F)	0.942	Ipin (M)	3.643	Ipin (M) & Mei Mei (F)	0.063
4.	Mail (M)	0.092	Ipin (M)	0.932	Mei Mei (F)	2.743	Mail (M)	0.059
5.	Opah (F)	0.082	Mail (M)	0.883	Mail (M)	1.893	Opah (F)	0.056
Movie								

1.	Lim (M), Badrol (M) & Ipin (M)	0.080	Ipin (M)	1.000	Badrol (M)	53.514	Lim (M) & Badrol (M)	0.030
2.	Upin (M) & Rajoo (M)	0.072	Upin (M)	0.938	Tok Dalang (M)	44.317	Rajoo (M)	0.028
3.	Tok Dalang (M), Muthu (M), Salleh (M) & Kak Ros (F)	0.065	Rajoo (M)	0.912	Lim (M)	41.995	Ipin (M)	0.026
4.	Pak Mail (M)	0.058	Kak Ros (F)	0.891	Pak Cik Bas (M)	41.000	Tok Dalang (M), Upin (M) & Muthu (M)	0.026
5.	Opah (F)	0.051	Lim (M)	0.888	Pak Mail (M)	37.452	Pak Mail (M) & Kak Ros (F)	0.025

2.3. ***Major, Minor and Extra Role Classification***

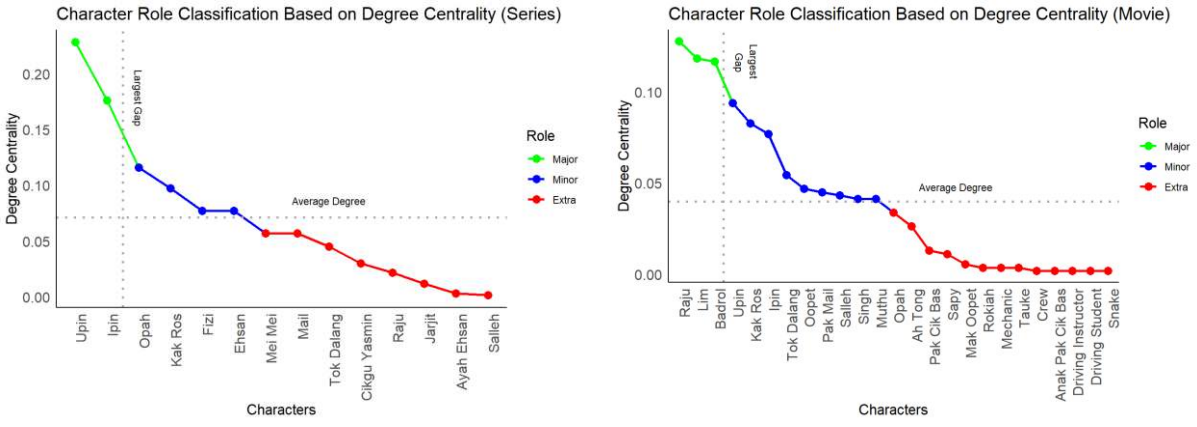
To classify characters into Major, Minor or Extra roles, the directed and weighted DC concept from the Methodology section is applied. Since the DC is weighted, characters with at least one interaction over a higher number of episodes (in the series) or five-minute intervals (in the movie) appear more central.

For the series in Fig. 3 (a), Upin and Ipin are the only two Major roles as the show revolves around their daily lives. Although Upin is much more talkative than Ipin, the largest gap in DC values is actually between Ipin and Opah. This highlights Opah as the most important Minor role, fitting her job as a supportive grandmother who constantly advises and guides her grandchildren. Kak Ros also plays a crucial part in shaping Upin and Ipin's upbringing. Additionally, Fizi and Ehsan are Minor roles due to their frequent interactions with the twins, whether playing together or discussing fasting habits. All remaining characters have DC values below 0.071 and are classified as Extras, meaning they have far fewer interactions. This is especially true for Salleh, whose only interaction occurs when Upin apologises for dirtying his porch in Episode 5 of Season 1 (Fig. 2). Surprisingly, Cikgu Yasmin is also an Extra, likely due to the limited number of kindergarten episodes, reducing her overall screen time.



Fig. 2 | Salleh's only interaction throughout the series is when Upin apologises to him in Episode 5 of Season 1

Considering the movie in Fig. 3 (b), two gaps share the largest value of 0.023, that is between Badrol and Upin, and the other between Kak Ros and Ipin. The first gap is chosen as the threshold to distinguish Major and Minor roles. As seen in Tab. 2, Rajoo, Lim and Badrol are Major roles due to their frequent interactions throughout the movie. Rajoo, being five years older than Upin and Ipin, converses more frequently with others, especially when strategising their escape from antagonists. Upin, Kak Ros, Ipin, Tok Dalang, Oopet, Pak Mail, Salleh, Singh and Muthu are Minor roles with directed and weighted DC values above 0.04. Evidently, while Kak Ros retains her Minor role, Opah is relegated to an Extra. This makes sense as Kak Ros actively protects the kids during their adventure, whereas Opah's role is significantly reduced to searching for the missing kids alongside other village elders. The three main antagonists, Pak Mail, Singh and Salleh are classified as Minor characters. Although they contribute tension and drama, their interactions are restricted to a specific subset of the main cast. Unlike Major roles that engage widely, these antagonists function as barriers to the story rather than serving as central figures. This reflects the producers' deliberate creative decisions in adapting character roles for different media formats.



(a) (b)

Fig. 3 | Role classification based on directed and weighted DC values for the (a) series and (b) movie respectively

2.4. *Null Configuration Model*

The null configuration model provides a benchmark for comparison with observed character network values, by preserving the original degree distribution while randomising or rewiring edges. This helps determine whether a character's prominence is authentic or simply due to chance, and whether degree distribution dictates character roles and prominence. Comparing observed centrality values to those from null models reveals whether a character's prominence stems from inherent structural properties beyond random connectivity patterns. Fig. 4 compares eigenvector, betweenness and closeness centralities between the null model and observed network for each character. Each plot illustrates mean centrality values from 1000 null model simulations, with error bars representing one standard deviation, alongside the observed values that indicate actual character centrality.

For the series, eigenvector and closeness centrality comparisons provide little insight, as all character values fall within one standard deviation of the null model. This implies that their prominence is driven by degree distribution and their central roles could be due to chance. However, betweenness centrality (BC) paints a compelling observation. While Upin and Ipin's observed BC values are lower than random expectation but still within the one standard deviation range, Ehsan stands out as a high outlier with significantly higher BC than expected.

Fig. 4 (d) shows that Anak Pak Cik Bas (bus driver's son), Mechanic, Pak Cik Bas, Pak Mail and Tok Dalang are low-eigenvector outliers, while Sapy (Rajoo's cow) is a high-eigenvector outlier in the movie. Peripheral characters like Anak Pak Cik Bas, Mechanic and Pak Cik Bas naturally exhibit low EC, but Pak Mail and Tok Dalang unexpectedly show lower connections than expected with key figures. Sapy's higher than expected EC is probably due to its strong bond with Rajoo, and its embedding within the most central network cluster. Even though Pak Cik Bas, Pak Mail and Tok Dalang are low-eigenvector outliers, Fig. 4 (e) portrays them as high-betweenness outliers. Pak Cik Bas connects his son and the mechanic, Pak Mail serves as a link to Tauke (a colloquial term for a Chinese business owner), while Tok Dalang connects the news crew, namely Rokiah and Crew, to the larger clusters. Being the main protagonists of the movie, Badrol and Lim exceed their expected BC values, reinforcing their significance in connecting various characters from different clusters. Conversely, Upin, Ipin and Muthu emerge as low-betweenness outliers, suggesting less dispersed interactions than expected.

Characters with lower-than-expected closeness centrality (CC) have wide-ranging null model values, indicating fluctuating positions in randomised networks. These include Anak Pak Cik Bas, Crew, Mechanic, Rokiah, Snake and Tauke. However, their observed CC values remain below the mean of those randomised fluctuations, reinforcing their roles as Extras in the movie as they occupy the network's periphery in Fig. 1 (b). Their limited accessibility is not merely a consequence of degree distribution, but rather a deliberate isolation from the network's core. While certain characters defy random assumptions in eigenvector and betweenness values, lower-than-expected CC signifies a peripheral or structurally distant character with minimal influence on the community.

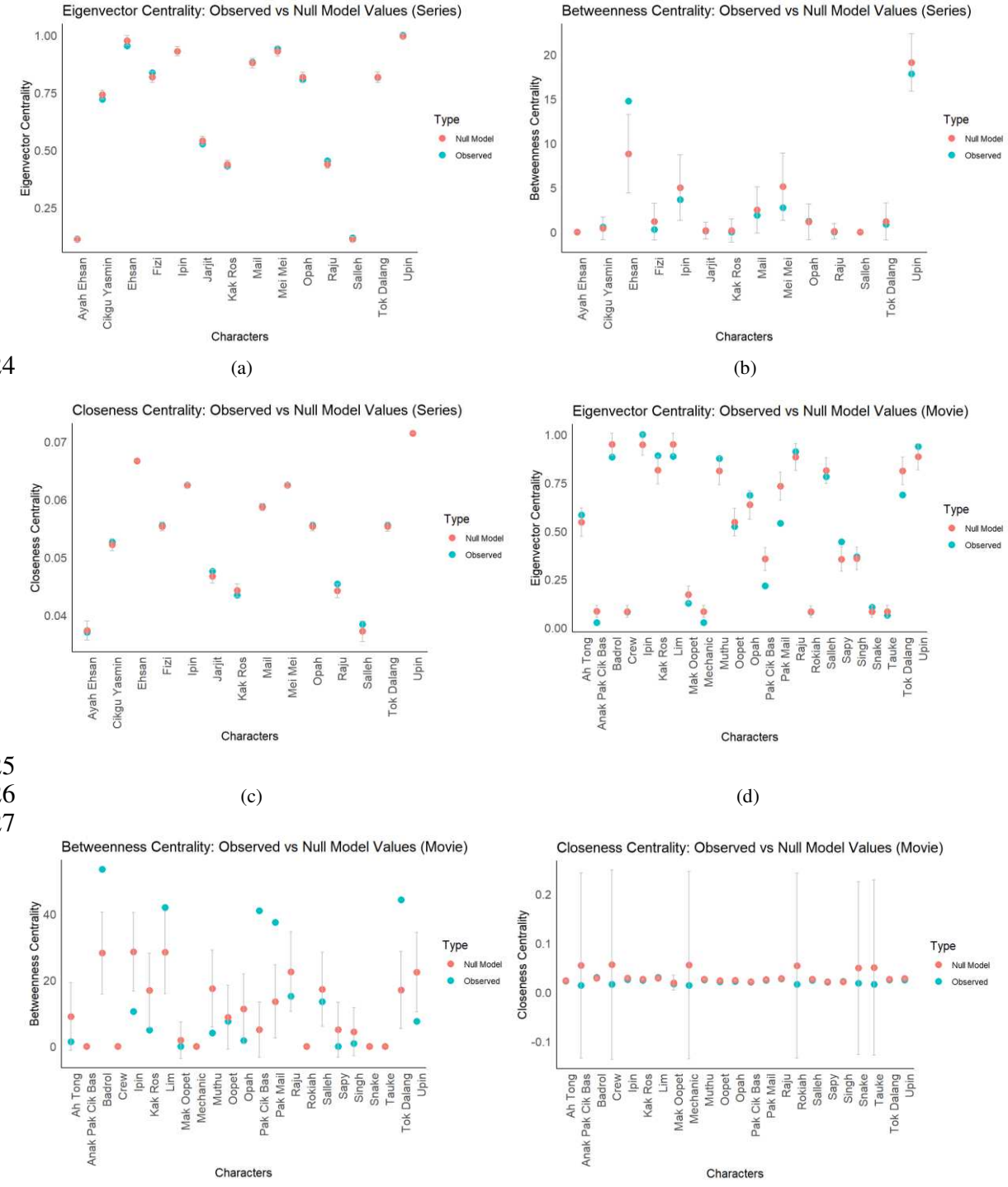


Fig. 4. | Observed versus null model eigenvector, betweenness and closeness centrality values for the (a, b, c) series and (d, e, f) movie networks

3. Discussion

SNA can help determine whether fictional networks mimic real-life social connectivity trends (Tan et al. 2014). Possessing small-world characteristics, the Upin & Ipin networks imitate real-world interactions to some extent, much like the Marvel comics network (Shi et al. 2021). This suggests that the show's social dynamics may be relatable to audiences as they reflect actual human interaction patterns, providing insight into how humans think and act in SEA, albeit some fantasy elements. Hence, animated character networks like Upin & Ipin could serve as proxies for real-life interactions, especially since actual contact data is challenging to obtain.

The character interactions in both the series and movie from Fig. 1 are fairly spread out, contrasting the rich-club effect in Marvel comics (Gleiser 2007), where highly connected characters tend to be more densely interconnected among themselves. Instead, the Upin & Ipin network is decentralised, where teamwork and communal relationships interplay. Unlike some protagonist-centered structures in Japanese manga (Sugishita & Masuda 2023), or the hero-centric structures common in Western media (Gleiser 2007), Upin & Ipin tends to adopt tightly knit communities. This resembles the team-based dynamics of X-men (Gleiser 2007), where collaboration takes precedence over the egocentric nature seen in narratives like Spider-Man or Captain America, in which heroes interact with most characters. The communal nature of Upin & Ipin mirrors real-life childhood dynamics, where kids naturally engage in group play and learn from one another. By prioritising collective interactions over individual heroism, Upin & Ipin embodies cultural values of camaraderie and harmony, distinguishing it from its more well-known global entertainment counterparts. The structural dynamics can inform filmmakers, scriptwriters and animators in crafting compelling SEA animations by producing relatable stories, even if they diverge from popular storytelling techniques.

An overview of Tab. 2 discloses a balance between races and genders in central roles, highlighting the producers' efforts to promote equality as it can strongly influence kids. Interestingly, centrality measures reflect key differences between the series and movie. In the movie, the prominence of Badrol, Lim and Rajoo marks a big narrative shift, focusing on broader tales within the Upin & Ipin universe with potential future spin-offs. Nevertheless, no spin-offs involving Badrol, Lim or Rajoo have materialised, reaffirming that the franchise remains anchored around the twins. Since the series predated the movie, the twins were already well-established, making the movie feel like a departure from Upin and Ipin rather than being the original broader vision. Rajoo's high directed and weighted DC likely stems from his unique ability to communicate with animals, a key fantasy-driven component of the movie that opposes the series' grounded themes. Meanwhile, Lim and Badrol's high centrality reflects their strong bond and recurring interactions with the villagers, which are crucial to the movie's plot. The shift from Upin and Ipin to newer characters emphasise the adaptability of animation towards evolving storylines while maintaining continuity within the Upin and Ipin universe.

Nonetheless, Upin and Ipin's high EC scores illustrate their deep embedding within the village's core social structure. Although Ipin is typically the quieter twin, he appears more outgoing in the movie, engaging in conversations with more characters, including key figures. This slight character change may have intended to enhance his appeal in the movie. Interestingly, Upin outperforms Ipin in directed and weighted DC, but Ipin ranks higher in undirected and unweighted degree. This indicates that Ipin converses with more characters, whereas Upin sustains conversations over multiple five-minute intervals. This distinction underscores the importance of understanding centrality measure interpretations when conducting analyses.

The relatively low centrality scores of main antagonists such as Pak Mail, Singh and Salleh suggest that their influence is localised rather than widespread, aligning with the show's nature.

Since media violence can increase aggression in children, minimising antagonist prominence may serve as a protective mechanism. A child's psychology is susceptible to powerful observational learning, internalising behaviours they witness (Huesmann 2007). Placing antagonists in "glorious" positions could encourage emulation, which is counterproductive. Thus media, especially animation, should strategically set protagonists in central positions, reinforcing moral messaging and upholding the principle that justice prevails.

The null configuration model reveals unique structural positioning of certain characters with prominence not solely dependent on degree distribution. In the movie, character centrality is driven by factors beyond degree distribution as observed values deviate more vigorously from null model expectations than in the series. This suggests that the movie's network structure is shaped by specific narrative choices with intentional story-driven connections that do not follow random social patterns. The strategically thought-out influential positions align with the movie's heightened interrelationship complexity.

Schauf and Escobar Varela (2018) examined Indonesia's centuries-old shadow puppet theatre called *wayang kulit* via null model ensembles based on degree distribution. They found that female characters exhibit high betweenness values despite limited appearances, revealing their social roles as bridges. These high betweenness outliers bridge peripheral characters that might otherwise be isolated from the giant component rather than being closely related to high-profile characters or traversing the network's core. Similarly, our findings highlight Ehsan as a high betweenness outlier due to his role in connecting his father, one of the two peripheral characters in the series (the other being Salleh), to the rest of the community. Without Ehsan, the only father figure of Seasons 1 and 2 would be entirely disconnected. Besides acting as a bridge, Ehsan's wealth and boastful nature initiates mild tensions, making him a critical figure in shaping both friendships and conflicts that add friction or depth to the story.

BC is a powerful metric in SNA, often associated with superspreaders (Abdul Razak & Expert 2021; Razak & Zamzuri 2021) who are individuals capable of transmitting something to many contacts. Regardless of low DC, their structural position renders them capable of controlling the flow of information, rumours, behaviours and diseases across different groups. In this study, characters with high BC, such as Ehsan in the series, or Tok Dalang in the movie, emerge as key spreaders of trends, conflicts or new behaviours. For example, Tok Dalang's rumour not only incites panic among villagers, but also reaches the newsreader, amplifying its circulation through the media. The animation network relates to real-life dynamics, where high BC individuals hold disproportionate influence. This makes them strategic targets for intervention in curbing misinformation and maintaining social harmony, or in accelerating the diffusion of essential messages.

Overall, by examining character interactions via a network perspective, series and movie narrative dynamics can be understood better. SNA offers an analytical and numerical approach to explain complex character relationships, systematically classifying characters based on their centrality and roles in driving the show. Character positioning shapes compelling stories, and a deep comprehension of these mechanics helps understand human interactions in the SEA region. Moreover, understanding narrative flows aids in replicating successful films that influence audience perception and identity. This study depicts how varying levels of complexity in character interactions help achieve distinct narrative goals within the same universe. The results serve as a foundational step towards comparative studies of animated series and movies using SNA. Future research could extend the methodology employed to other animations or incorporate dynamic network modelling to track relationship evolution.

4. Data and Methodology

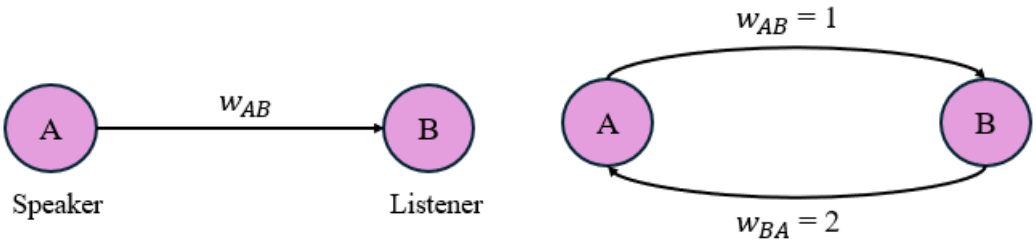
4.1. *Data*

The data collection stage is explained in terms of the identification and characterisation of character interactions, focusing on relationships to gain insights into character structural prominence. Character interaction data can be collected through several ways, with no definitive general approach due to inherent subjectivity. Nonetheless, the data sampling method in this study was pre-specified before being conducted. The resulting character conversational network is constructed based on verbal interactions or dialogues between two characters, incorporating speaker and listener identification. The character network is a graph, denoted by $G = (V, E)$, in which V is the set of nodes or vertices that represent each character, whereas E is the set of edges representing the dialogues taking place between two characters. Speaker-listener identification ensures directed networks as illustrated in the conversational graph in Fig. 5 (a).

The graph in Fig. 5 (a) consists of two characters, where node A is the speaker (the edge originates from it), while node B is the listener (the edge points towards it). The character network is both directed and weighted, with the weight denoted as w_{AB}. To determine edge weights, we adopt a mix between the Boolean and numerical score approach to represent the intensity or frequency of interactions. For each episode in the series, a directed edge from node A to node B takes one of two possible values:

$$w_{AB} = \begin{cases} 1, & \text{if } A \text{ speaks to } B \text{ at least once} \\ 0, & \text{if } A \text{ does not speak to } B \text{ at all} \end{cases}$$

The weights are accumulated over all 18 episodes, resulting in a maximum possible weight of 18 for any directed edge. For instance, in Episode 1, node A speaks to node B once, and vice versa. In Episode 2, node B speaks to node A three times (recorded as a weight of 1 following episode-based weighting), while node A does not speak to node B. The two episodes lead to a directed edge from node A to node B with weight 1, and a directed edge from node B to node A with weight 2, as shown in Fig. 5 (b).



(a) (b)

Fig. 5 | (a) A directed and weighted conversational network that requires speaker and listener identification, and (b) an example of a simplified network between two characters over two episodes.

The same procedure applies to the movie, but instead of episodes, interactions are recorded in five-minute intervals. The 93-minute and 58-second movie is divided into 18 intervals (with the final interval lasting 3 minutes and 58 seconds), leading to a maximum possible weight of 18 for any directed edge, similar to the series. Since the series is also made up of 18 episodes with a duration of 4-5 minutes each, this approach enhances comparability between the two

media formats. Without access to the show's scripts, character interactions were manually extracted by watching the show. Employing edge weights based on dialog duration would be impractical, thus we opt instead for a simplified Boolean-numerical scoring approach. This unique manual method is well-suited for short series or movies, offering flexibility in establishing custom rules and ensuring easy validation, paving the way for character network studies.

The following rules were applied when recording interactions:

- Only verbal interactions are recorded. Non-verbal actions such as slapping, hugging, nodding and so on are excluded.
- The speaker must address only one person at a time, unless explicitly naming multiple recipients before speaking to them. For example, an interaction is not recorded if Kak Ros speaks to both Upin and Ipin at the same time without naming them, since it is not clear who she is particularly addressing. However, if Kak Ros explicitly says, "Upin, Ipin, dinner is ready", edges are generated from Kak Ros to Upin, and Kak Ros to Ipin, both having weights equal to 1.
- Replies count as interactions. For instance, a weight of 1 is assigned from node *A* to node *B* and vice versa if *A* asks a question and *B* responds.

  Upin: Ipin, do you know how to answer this question?
  Ipin: Yes, it is easy.
- Interactions are excluded if the addressee is unclear, the speaker is off-screen and unidentifiable, or if self-talk/internal monologues occur (no self-loops in the networks).
- Filler words such as *hm* and *mm*, or sighs are not recorded.

Once data collection is completed, character networks are generated for the series and movie respectively. Further analyses are performed on the networks via RStudio (version 4.2.2), particularly leveraging the *igraph* and *sna* libraries.

4.2. *Methodology*

Four major analyses are carried out on the networks. While some analyses make use of the directed and weighted network versions, others are performed on the undirected and unweighted versions to ensure transparency and clarity. Similarly, Rahim et al. (2023) applied simplified undirected-unweighted forms for betweenness centrality (BC) computations in their student network study, as the formulas are more well-defined in literature.

4.2.1. *Community Detection and Global Network Statistics*

Since a complex web of interactions can be described by common characteristics such as community structure (Barabási & Pósfai 2016; Newman 2010), the networks in this study are visualised alongside their detected communities. Community detection identifies groups or clusters within a network, in which nodes belonging to the same cluster are more densely connected than those from different clusters. There are several community detection algorithms, each with its own interpretation. To select a suitable algorithm, the modularity of a community structure is computed, where higher scores indicate better partitioning of nodes. In this study, the Walktrap algorithm is selected for its ability to achieve high modularity and logical node clusters. The Walktrap algorithm as implemented by Pons and Latapy (2005), leverages the principle that short random walks tend to remain within the same cluster. The cluster_walktrap function in *igraph* effectively captures community structures while being computationally efficient.

Once the networks are constructed, global network statistics are calculated to further analyse the overall structure of the series and movie's undirected and unweighted networks.

- The global clustering coefficient (GCC), or transitivity as it is called in R, calculates the total number of closed triangles in the whole character network as shown in Eq. (1). R ignores the edge direction for directed graphs, rendering the network undirected. There are several approaches to computing the GCC, which was first defined by Watts and Strogatz (1998). However, the most common way to define it as presented by Newman (2010), is as follows:

$$\text{Global Clustering Coefficient} = \frac{3\times(\text{number of triangles})}{\text{number of connected triplets}} \quad (1)$$

- The shortest path, also known as the geodesic, is the path between the node pair with the smallest number of edges. The mean_distance function in *igraph* obtains the average path length of the whole network by calculating the shortest paths between all node pairs using an unweighted breadth-first search algorithm.
- Network density represents the actual number of edges observed in the network over the maximum number of edges possible, computed by the edge_density function in *igraph*. Essentially, the probability that a randomly picked pair of nodes is connected is equal to the density of the network.

$$\text{Density} = \frac{\text{actual number of edges}}{\frac{N(N-1)}{2}} \quad (2)$$

where N is the total number of nodes in the network.

4.2.2. *Centrality Measures*

One of the primary analyses in SNA is determining the most influential or important nodes within a network (Wasserman & Faust 1994). To further investigate how characters are embedded within the network, centrality measures help identify key characters on the show based on their relationships and network positions. Some of the most well-known and regularly used centrality measures include degree, eigenvector, betweenness and closeness centralities. Each centrality measure has its own definition and algorithm, resulting in the possibility of different central characters for each measure. These central characters hold specific roles and strategic locations in the network depending on the type of centrality applied.

Degree centrality (DC) is one of the simplest centrality metrics due to its intuitive definition, which considers the local neighbourhood of a node, that is, a node's direct neighbours. DC looks at the degree of connections a given node has with others. In a directed and weighted network, the weight and direction of interactions (incoming and/or outgoing) are taken into account. The directed and weighted DC for node *i* is computed as follows, applying the method used in Park et al.'s (2012) study and implemented in *igraph*:

$$DC_i(\text{directed and weighted}) = \frac{\left(\sum_{j=1}^{\alpha} w_{ji} + \sum_{k=1}^{\beta} w_{ik}\right)}{2\times\sum_{j=1}^{N}\sum_{k=1}^{N} w_{jk}} \quad (3)$$

where α is the number of incoming connections and β is the number of outgoing connections. The numerator of Eq. (3) represents the sum of all incoming and outgoing edge weights for node i, while the denominator represents the summation of all edge weights present in the

network. This formula provides a normalised degree score, which allows for fair comparisons across networks of different sizes.

The undirected and unweighted version of DC simply counts the number of connections a node has (Freeman 1978). Since all connections are assumed to be of equal importance, nodes with a higher number of connections are deemed more influential. The rescaled formula (to sum to 1) for the undirected and unweighted DC is:

$$DC_i(\text{undirected and unweighted}) = \frac{\sum_{j=1}^{N} A_{ij}}{2\times\text{number of edges}} \tag{4}$$

where $A_{ij} = 1$ if characters i and j interact, and 0 otherwise. Here, A denotes the adjacency matrix containing the interactions between all node pairs. Notably, directed and weighted DC measures the strength of connections, whereas undirected and unweighted DC considers the number of connections. From this point forward, all centrality measures are computed on the undirected and unweighted networks.

Unlike DC, which treats all edges equally, eigenvector centrality (EC) is a natural extension to DC that measures the importance of a node based on its relationship with other influential nodes (Bonacich 1987). Therefore, characters that interact with other influential characters may have higher EC scores as compared to those connected to less central characters. EC is calculated as follows:

$$EC_i = \frac{1}{\lambda}\sum_{j=1}^{N} A_{ij}x_j \tag{5}$$

with λ denoting the largest eigenvalue of A, while x_j is the leading eigenvector of character j.

Besides that, BC emphasises nodes that frequently lie on the shortest paths between other node pairs. BC can determine critical nodes that serve as bridges between different parts of the network, thereby influencing the flow of information. A node has a high BC score if many other nodes need to traverse it to reach one another. BC can be defined by Eq. (6) (Freeman 1977).

$$BC_i = \sum_{h<k}^{N} \frac{b_{hk}(i)}{b_{hk}}, h \neq k \neq i \tag{6}$$

where b_{hk} denotes the number of shortest paths that connect characters h and k. More specifically, $b_{hk}(i)$ represents the number of those shortest paths that traverse character i.

Closeness centrality (CC) quantifies the proximity of each character to every other character in the network. The CC of character i is defined as the average shortest path from character i to all the other characters. A character with high CC is more efficient in spreading information throughout the network due to its proximity to most of the others. CC can be mathematically expressed as shown below (Freeman 1978):

$$CC_i = \left[\sum_{j=1}^{N} d(i,j)\right]^{-1} \tag{7}$$

with $d(i,j)$ representing the number of edges in the shortest path connecting characters i and j.

Rather than focusing on immediate neighbours, BC and CC are higher-order centrality measures that capture a character's placement in the network amid long chains of interaction links. Note that the EC, BC and CC of the movie's network are only computed on the giant component, seeing as the network is disconnected (refer to Fig. 1 (b)). A disconnected graph is a network that is made up of several separate components with no paths connecting them,

whereas in a connected graph, every single node can reach one another via a path. Restricting EC, BC and CC computations to the giant component (largest connected subgraph) enables fair comparisons across these three centrality measures, as EC and CC are only meaningful for connected graphs.

4.2.3. *Major, Minor and Extra Role Classification*

Characters are classified into one of three roles, namely a Major, Minor or Extra. Role classifications provide deeper insight into characters who drive the plot or contribute to the story's development, separating them from background characters. Comparing character roles between the series and movie also helps to reveal shifts in character prominence. To determine roles, the directed and weighted DC from Eq. (3) is averaged as follows:

$$\text{Average DC} = \frac{\sum_{i=1}^{N} DC_i}{N} \tag{8}$$

General definitions of each role type (Park et al. 2012):

- Major: Leading characters that interact with many others, have constant appearances and significantly impact the show and lead story development.
- Minor: Support Major characters, besides supporting or even leading story development.
- Extra: Appear briefly, often in as few as one or two scenes.

Role classification procedure:

(1) Calculate the DC for each character using the directed and weighted version in Eq. (3).
(2) Rank characters according to their DC values in descending order.
(3) Calculate the average DC value.
(4) Compute the difference in DC values between each consecutive pair of characters.
(5) Determine the largest gap between consecutive DC values from step 4, which serves as a threshold for classifying Major and Minor roles.

Role classification is executed based on the following criteria:

- Major: Characters with DC values exceeding both the average DC and the largest gap determined in step 5. Technically, there will be two characters with the largest gap in DC values. The upper character and all those characters with DC values higher than it have Major roles.
- Minor: Characters with DC values exceeding the average DC, but below the largest gap determined in step 5. The lower character in this gap, and all those characters with DC values lower than it (but still above the average DC) have Minor roles.
- Extra: Characters with DC values below the average DC.

This role classification method was utilised by Weng et al. (2009), who categorised characters into two classes, namely Leading and Supporting roles. Park et al. (2012) then extended this classification by introducing an additonal role, called Extra, which we employ in this study.

4.2.4. *Null Configuration Model*

A null model is a simulated network used as a baseline to assess whether observed patterns in the original network are statistically significant or merely expected outcomes generated by chance. Null models preserve specific features of interest in the original network, while randomising other aspects. The configuration model is a commonly used null model that preserves the degree sequence of the original network in its randomly simulated networks. This

model was introduced by Bollobás (1980), although the concept was based on earlier work by Bender and Canfield (1978). In this study, the null configuration model evaluates whether a character is truly central or if their centrality is merely a byproduct of their degree distribution, occuring by chance. The following steps explain how the null configuration model analysis is carried out.

(1) Each character in the original series or movie network retains its degree. Using the provided degree sequence, `sample_degseq` generates random networks, where the `method` argument can be set to prevent multiple edges or self-loops in the simulations.

(2) Edges are rewired to form random connections. For instance, consider a network with five nodes, namely node *A*, *B*, *C*, *D* and *E*, with the degree sequence {3, 1, 2, 1, 1} as shown in Figure 6 (a). Envision each of the five nodes possessing half-edges, called stubs, corresponding to their degree values as shown in Figure 6 (b). Every stub is randomly paired with another stub to form a full edge, ensuring that the simulated network has the exact same degree distribution as the original network. Figure 6 (c) and (d) depict two examples of random networks that can be generated by the null configuration model.

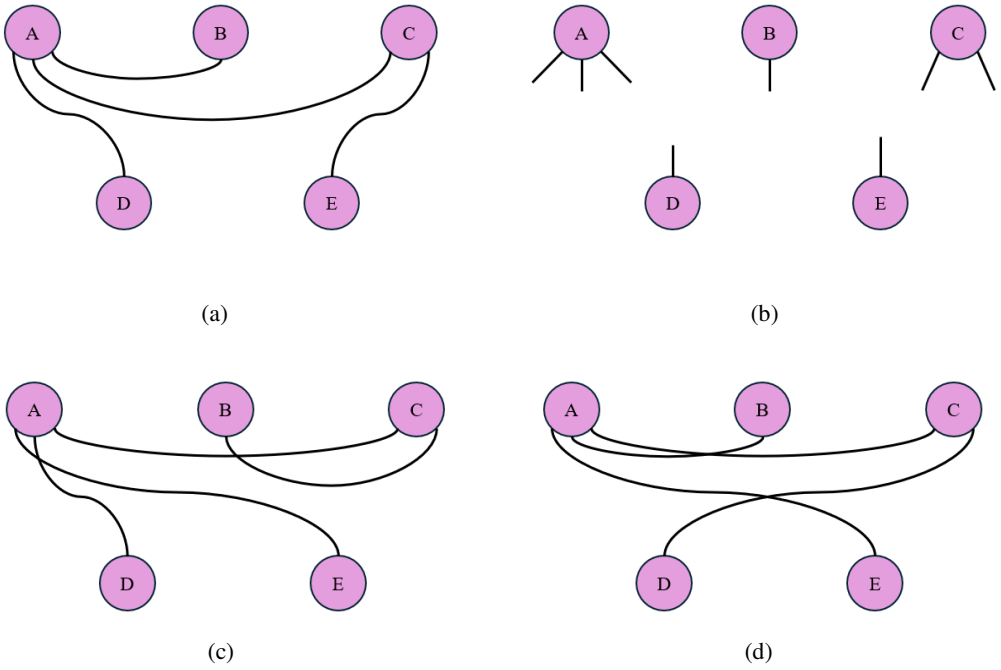


Fig. 6 | Example of a (a) network, the (b) stubs for each of its nodes, and (c, d) two realisations of the null configuration model

(3) The generation of random networks is repeated 1000 times to obtain statistically robust results.

(4) EC, BC and CC values of characters in each of the simulated networks are computed (restricted to the giant component in the movie network).

(5) Mean centrality values across the ensemble of simulated networks are computed. These mean values represent the expected centrality of each character as generated by the null configuration model.

(6) Centrality values from the null model and original network are visualised in a single plot (one plot for each centrality measure) to facilitate comparisons between their values.
(7) Error bars of one standard deviation are included to identify potential outliers. Characters with observed centrality values that fall outside the one standard deviation range are outliers, holding a structurally significant position that is not random or merely a result of degree distribution. Conversely, characters whose observed values closely match the null model likely owe their centrality to degree constraints, emerging by chance.

References

Abdul Razak, F., Ahmad Shahabuddin, F. & Sarah Nik Zamri, N. 2019. Analyzing research collaborations within the School of Mathematical Sciences, UKM using Graph Theory. *Journal of Physics: Conference Series* 1212(1): 012033.

Abdul Razak, F. & Expert, P. 2021. Modelling the spread of COVID-19 on Malaysian contact networks for practical reopening strategies in an institutional setting. *Sains Malaysiana* 50(5): 1497–1509.

Alberich, R., Miro-Julia, J. & Rossello, F. 2002. Marvel Universe looks almost like a real social network.

ASEAN Intellectual Property Portal. n.d. Animating an international brand.

Azizul Rahman Ismail. 2025. QuickCheck: Is there an official Malaysian Ultraman? https://www.thestar.com.my/news/true-or-not/2025/01/23/quickcheck-is-there-an-official-malaysian-ultraman (13 March 2025).

Bahari, N.N., Expert, P. & Razak, F.A. 2023. An analysis of actors in Malay films: Small worlds, centralities and genre diversity. *Mathematics* 11(5): 1252.

Barabási, A.-L. & Pósfai, M. 2016. *Network Science*. Cambridge: Cambridge University Press.

Bender, E.A. & Canfield, E.R. 1978. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A* 24(3): 296–307.

Biddle, B.J. 1986. Recent developments in role theory. *Annual Review of Sociology* 12(1): 67–92.

Bollobás, B. 1980. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics* 1(4): 311–316.

Bonacich, P. 1987. Power and centrality: A family of measures. *American Journal of Sociology* 92(5): 1170–1182.

Bost, X., Gueye, S., Labatut, V., Larson, M., Linarès, G., Malinas, D. & Roth, R. 2019. Remembering winter was coming. *Multimedia Tools and Applications* 78(24): 35373–35399.

Bost, X., Labatut, V., Gueye, S. & Linarès, G. 2016. Narrative smoothing: Dynamic conversational network for the analysis of TV series plots. In *2nd International Workshop on Dynamics in Networks, in Conjunction with the 2016 IEEE/ACM International Conference ASONAM*. pp. 1111–1118. IEEE.

Chowdhury, T., Muhuri, S., Chakraborty, S. & Chakraborty, S.N. 2019. Analysis of adapted films and stories based on social network. *IEEE Transactions on Computational Social Systems* 6(5): 858–869.

Fandom. n.d.-a. Geng: Pengembaraan Bermula. https://upinipin.fandom.com/ms/wiki/Geng:_Pengembaraan_Bermula (15 March 2025).

Fandom. n.d.-b. Upin & Ipin. https://upinipin.fandom.com/wiki/Upin_%26_Ipin#Episodes (2 February 2025).
Fonseca, J.D., Ferreira, H.M., Silva, M.C. da & Lima, I.V. 2023. The animation “Purl”: An analysis from the perspective of Sociointeractional Semiotics. *Texto Livre* 16: e45482.
Freeman, L.C. 1977. A set of measures of centrality based on betweenness. *Sociometry* 40(1): 35–41.
Freeman, L.C. 1978. Centrality in social networks conceptual clarification. *Social Networks* 1(3): 215–239.
Giri, S., Chaudhary, S. & Gautam, B. 2024. Analyzing social networks of actors in movies and TV shows.
Gleiser, P.M. 2007. How to become a superhero. *Journal of Statistical Mechanics: Theory and Experiment* (9): P09020.
Huesmann, L.R. 2007. The impact of electronic media violence: Scientific theory and research. *Journal of Adolescent Health* 41(6): S6–S13.
Jung, J.J., You, E. & Park, S.-B. 2013. Emotion-based character clustering for managing story-based contents: A cinemetric analysis. *Multimedia Tools and Applications* 65(1): 29–45.
Karinthy, F. 1929. *Chains*.
Labatut, V. & Bost, X. 2019. Extraction and analysis of fictional character networks. *ACM Computing Surveys* 52(5): 1–40.
Lee, E., Karimi, F., Wagner, C., Jo, H.-H., Strohmaier, M. & Galesic, M. 2019. Homophily and minority-group size explain perception biases in social networks. *Nature Human Behaviour* 3(10): 1078–1087.
Lee, J. & Yeung, C.Y. 2012. Extracting networks of people and places from literary texts. In *Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation*. pp. 209–218.
Mandić, M., Škobić, D. & Martinović, G. 2018. Clique comparison and homophily detection in telecom social networks. *International Journal of Electrical and Computer Engineering Systems* 9(2): 81–87.
Milgram, S. 1967. The small-world problem. *Psychology Today* 1(1): 61–67.
Min, S. & Park, J. 2016. Network science and narratives: Basic model and application to Victor Hugo’s Les Misérables. In Cherifi, H., Gonçalves, B., Menezes, R., & Sinatra, R. (eds.). *Complex Networks VII*. Vol. 644 pp. 257–265. Springer, Cham.
Murakami, H., Nagaoka, Y. & Kyogoku, R. 2020. Creating character networks with kinship relations from comics. *International Journal of Service and Knowledge Management* 4(1): 1–26.
Nan, C.-J., Kim, K.-M. & Zhang, B.-T. 2015. Social network analysis of TV drama characters via deep concept hierarchies. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. pp. 831–836. New York, NY, USA: ACM.
Newman, M.E.J. 2010. *Networks : An Introduction*. 1st Ed. New York: Oxford University Press.
Park, S.-B., Oh, K.-J. & Jo, G.-S. 2012. Social network analysis in a movie using character-net. *Multimedia Tools and Applications* 59(2): 601–627.
Pons, P. & Latapy, M. 2005. Computing communities in large networks using random walks.
Prado, S.D., Dahmen, S.R., Bazzan, A.L.C., Carron, P. Mac & Kenna, R. 2016. Temporal network analysis of literary texts. *Advances in Complex Systems* 19(3): 1650005.

Rahim, N.Z.A., Bahari, N.N., Azzimi, N.S.M., Zamzuri, Z.H., Bahaludin, H., Mohammad, N.F. & Razak, F.A. 2023. Comparing friends and peer tutors amidst COVID-19 using social network analysis. *Mathematics* 11(4): 1053.

Ratanarajah, J., Razak, F.A. & Zamzuri, Z.H. 2020. Peer tutor network and academic performance: A UKM pilot study. In *AIP Conference Proceedings*. Vol. 2266 p. 050013. American Institute of Physics Inc.

Razak, F.A. & Zamzuri, Z.H. 2021. Modelling heterogeneity and super spreaders of the COVID-19 spread through Malaysian networks. *Symmetry* 13(10): 1954.

Ribka Yulista Devi Purnamasari. 2020. Pengaruh film kartun Upin Ipin terhadap perilaku anak usia 7-9 tahun. *Jurnal Edukasi Nonformal* 121–126.

Schauf, A. & Escobar Varela, M. 2018. Searching for hidden bridges in co-occurrence networks from Javanese wayang kulit. *Journal of Historical Network Research* 2(1): 26–52.

Shi, C., Yu, X. & Ren, Z. 2021. How the Avengers assembled? Analysis of Marvel hero social network.

Sugishita, K. & Masuda, N. 2023. Social network analysis of manga: Similarities to real-world social networks and trends over decades. *Applied Network Science* 8(1): 79.

Tan, M.S.A., Ujum, E.A. & Ratnavelu, K. 2014. A character network study of two Sci-Fi TV series. In *AIP Conference Proceedings*. Vol. 1588 pp. 246–251. American Institute of Physics Inc.

Wasserman, S. & Faust, K. 1994. *Social Network Analysis: Methods and Applications*. 1st Ed. Vol. 8. Cambridge: Cambridge University Press.

Watts, D. & Strogatz, S. 1998. Collective dynamics of 'small-world' networks. *Nature* 393: 440–442.

Weiringen, S. 2024. Double the entertainment: Open-world 'Upin & Ipin' game headed to consoles. https://www.thestar.com.my/tech/tech-news/2024/07/06/double-the-entertainment-open-world-upin--ipin-game-headed-to-consoles (18 March 2025).

Weng, C.-Y., Chu, W.-T. & Wu, J.-L. 2009. RoleNet: Movie analysis from the perspective of social networks. *IEEE Transactions on Multimedia* 11(2): 256–271.

Yeh, M.-C. & Wu, W.-P. 2014. Clustering faces in movies using an automatically constructed social network. *IEEE MultiMedia* 21(2): 22–31.

Yulianeta & Amandangi, D.P. 2021. Upin-Ipin: Gender issues in early childhood education. In *Proceedings of the First Transnational Webinar on Adult and Continuing Education (TRACED 2020)*. pp. 166–169.

Zaefarian, G., Misra, S., Koval, M. & Iurkov, V. 2022. Editorial: Social network analysis in marketing: A step-by-step guide for researchers. *Industrial Marketing Management* 107: A11–A24.

Zeanaaima, M.Y. 2022. Upin dan Ipin tembusi UK, Australia, 'cakap' Hindustan di TV India, kata pencipta animasi. https://www.freemalaysiatoday.com/category/bahasa/tempatan/2022/09/17/upin-dan-ipin-tembusi-uk-australia-cakap-hindustan-di-tv-india-kata-pencipta-animasi/ (2 February 2025).

Competing Interests

The authors declare no competing interests.

Data Availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Ethical Approval Statement

This article does not contain any studies with human participants performed by any of the authors. Therefore, no ethics approval was required.

Informed Consent Statement

As this study does not involve human participants or personal data, informed consent was not applicable.

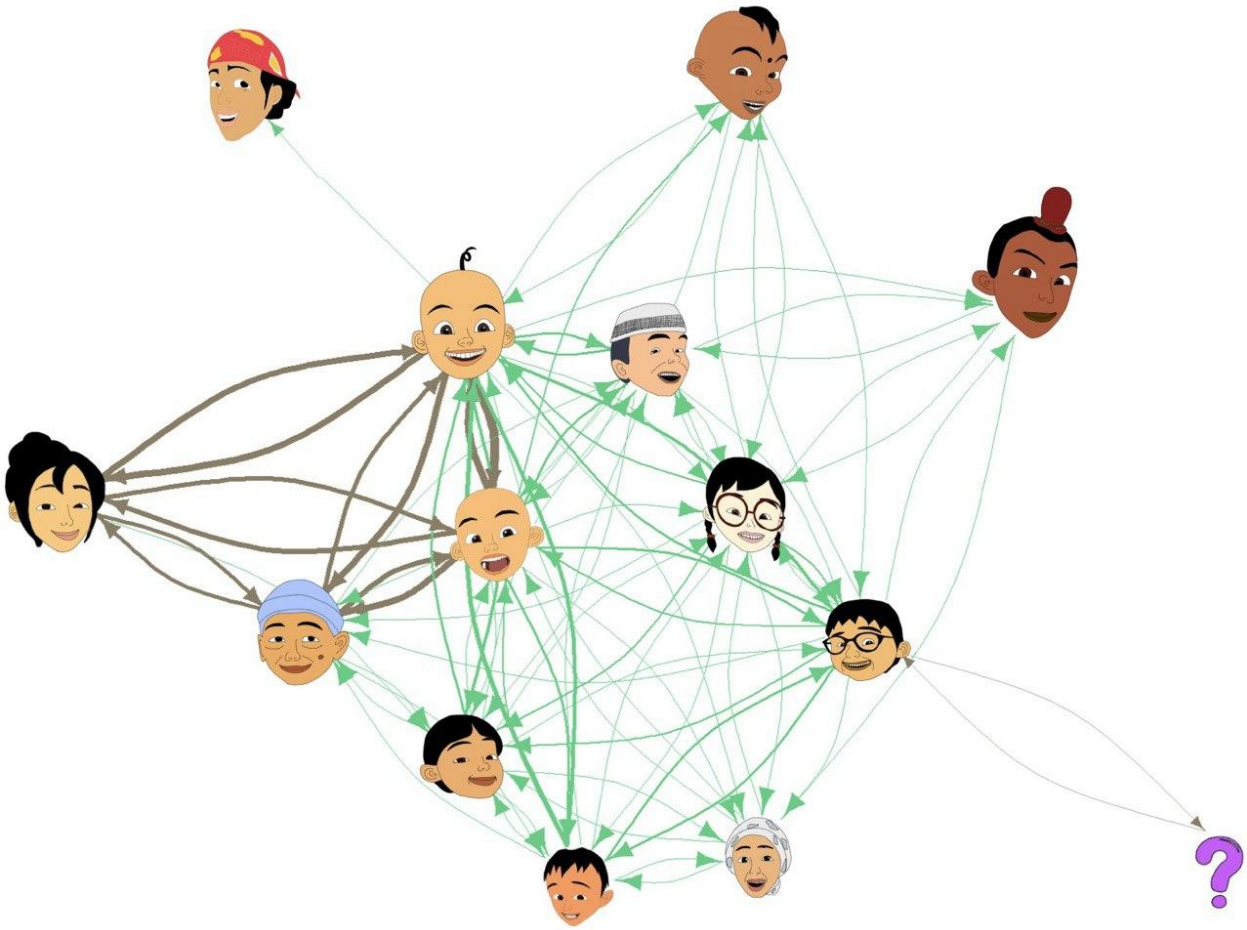
Author Contributions

Conceptualisation: [Author 1], [Author 2]; data curation: [Author 1]; formal analysis and investigation: [Author 1]; funding acquisition: [Author 2]; methodology: [Author 1], [Author 2]; project administration: [Author 1], [Author 2]; resources: [Author 1], [Author 2]; software: [Author 1]; supervision: [Author 2]; validation: [Author 2]; visualisation: [Author 1]; writing—original draft: [Author 1]; writing—review and editing: [Author 2], [Author 1]. All authors read and approved of the final manuscript.

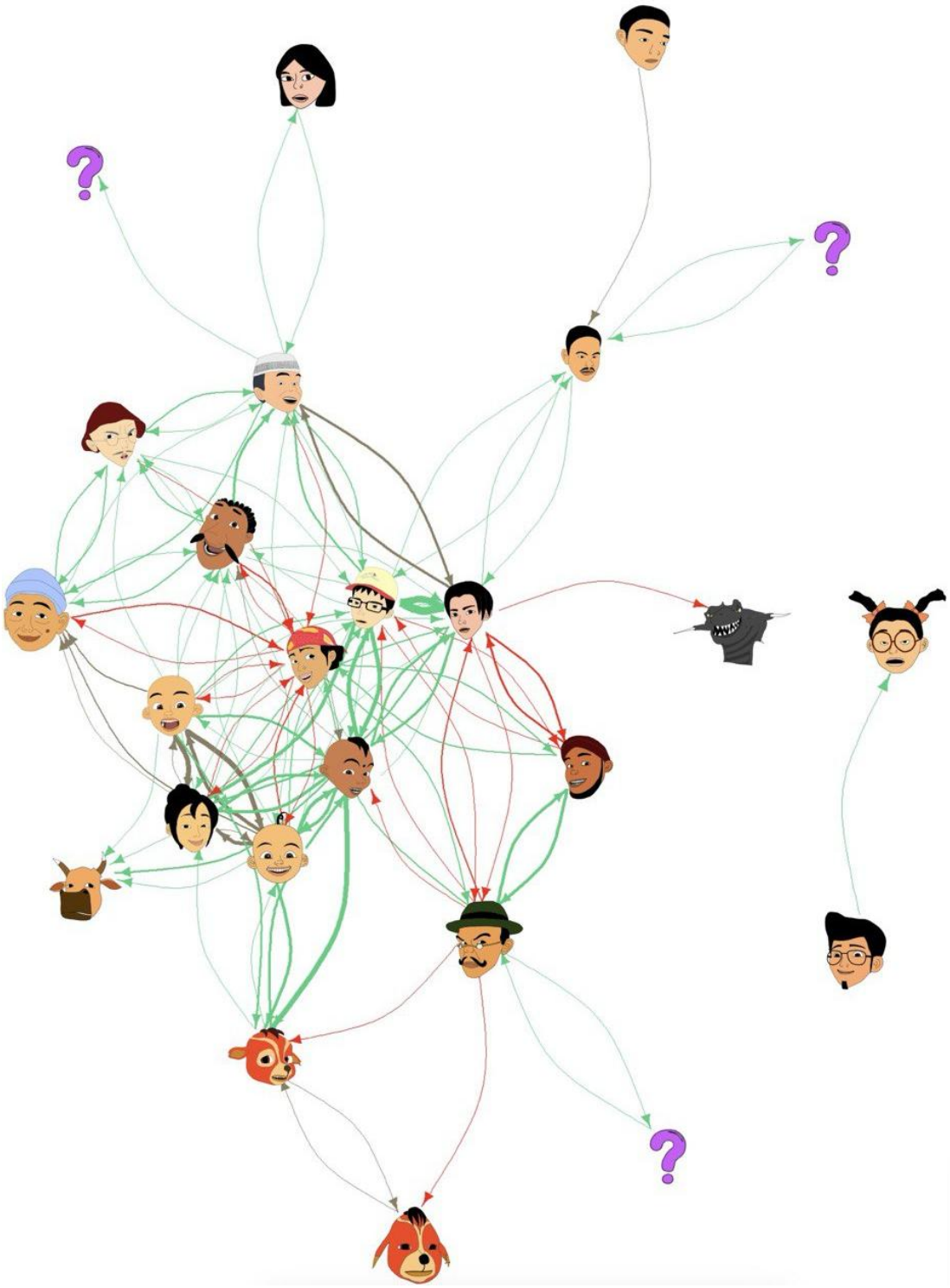
AI Disclosure

The authors used OpenAI's ChatGPT during the manuscript preparation stage to assist in language polishing. All intellectual contributions, data collection, analyses, interpretations and write-ups were performed by the authors, with the final draft being reviewed and approved by all the authors.

838 **Extended Data**



839 (a)



(b)

Extended Data Fig. 1 | Alternative versions of Fig. 1 (a) and (b) depict network nodes using actual character images instead of conventional shapes such as circles, squares or triangles. A question mark represents characters who only appear off-screen in the series and movie analysed, as their appearances remain unknown at this point of the study.

845 Extended Data Tab. 1 | Characters, their appearances, roles and/or personalities

Character	Gender	Race / Species	Appearance and/or Role and/or Personality
Upin	Male	Malay	The older twin (five years old), distinguished by a single strand of hair on his bald head and is often seen wearing a yellow tank top with the letter "U" printed on it. Upin is cheerful and talks a lot more as compared to Ipin.
Ipin	Male	Malay	The younger twin, completely bald and wears a blue tank top with the letter "I" printed on it, along with a little red scarf tied around his neck. Ipin is the shy and quieter twin who loves repeating the same word thrice when agreeing with someone. His tagline is "*Betul, Betul, Betul"*, which translates to "Correct, Correct, Correct".
Opah (grandmother)	Female	Malay	Her real name is Mak Uda, but she is addressed as Opah as a sign of respect by her grandchildren (Upin, Ipin, Kak Ros) and their friends. She is gentle, loving and often spoils her grandsons. Being the head of the family, she is knowledgeable about religion and life.
Kak Ros (Sister Ros)	Female	Malay	Upin and Ipin's older sister, has typical SEA features such as black hair, brown eyes and tan skin. Known to be strict as she often reprimands her younger brothers when they misbehave. Nevertheless, she deeply cares for them.
Tok Dalang	Male	Malay	Headman of the village and shadow puppet champion.
Ehsan	Male	Malay	A chubby boy with thick spectacles, serving as the class monitor for the friend group at kindergarten. He comes from a wealthy family and loves to show off his belongings, making him appear boastful.
Ayah Ehsan (Ehsan's father)	Male	Malay	Briefly appears off-screen in Episode 15 of Season 2. He affectionately calls Ehsan "*Intan payung*", a term of endearment similar to darling, sweetheart, or the apple of my eye.
Fizi	Male	Malay	Often seen with Ehsan and is known for his habit of laughing at his friends and making unintentional remarks.
Mail	Male	Malay	A calm and independent boy who is a part-time business entrepreneur. He is famous for selling fried chicken with the tagline, "*Dua seringgit*", meaning "Two for one ringgit (Malaysian currency)".
Mei Mei	Female	Chinese	A sweet, kind and caring girl who often advises her friends. She has fair skin, wears spectacles and usually ties her hair in pigtails.
Jarjit	Male	Sikh	A comedic character who loves reciting spontaneous two stanza poems starting with "*Dua tiga*", which translates to "Two three". He wears a turban, which is culturally worn by Sikh males.
Cikgu Yasmin (Teacher Yasmin)	Female	Malay	Kindergarten teacher at Mesra Kindergarten. She is often seen giving advice to the kids during the holy month of Ramadan.
Rajoo	Male	Indian	The oldest in the friend group (10 years old), twice the other kids' age. He has the unique ability to communicate with animals and is often seen with his cow, Sapy. He wears a bindi as an indication of his Hindu faith.
Muthu	Male	Indian	The village's most well-known hawker, known for selling a variety of food at his stall. He is Rajoo's father.
Salleh	Male	Malay	Known for his softer, more feminine personality. He prefers to call himself Sally instead of Salleh and is easily irritated.

Badrol	Male	Malay	Tok Dalang's grandson from Kuala Lumpur, who usually visits the village during festivities. He is a main character in the movie, visiting the village with his best friend, Lim, after seeing his grandfather on the news talking about the Durian Ghost. He is an energetic person who joins the kids on their adventure, but is eventually captured by the antagonists alongside Lim. He wears a jersey throughout the movie, reflecting his active nature.
Lim	Male	Chinese	Badrol's best friend who accompanies him to Kampung Durian Runtuh. He is a highly prepared person, but often overpacks unnecessary items, causing his backpack to be super heavy. He wears a "Tourism Malaysia" cap throughout the movie.
Ah Tong	Male	Chinese	A veteran that is often seen wearing a red jacket and round spectacles. He is brave and carries a rifle in the movie.
Sapy	Female	Animal (Cow)	Rajoo's pet cow.
Oopet	Male	Unknown creature	A mysterious creature that enters the village from another world through a portal hidden in a cave. He secretly eats durians in the village, leading villagers to believe in a "Durian Ghost". Oopet seeks to save his mother, who was taken by poachers (Pak Mail, Singh, Salleh) before they can return home. Only Rajoo can understand Oopet, making him the mediator between Oopet and others in the movie.
Mak Oopet (Oopet's mother)	Female	Unknown creature	Very protective of her son, fiercely defending him. She breaks free from captivity and chases down the poachers when she sees them hurting him.
Pak Mail	Male	Malay	The main antagonist in the movie, leading the poaching activities alongside Singh and Salleh. He is quite fierce and irritable. Pak Mail and Singh escape into Oopet's world when they are chased by Oopet's mother, but are most probably killed or injured there.
Singh	Male	Sikh	Pak Mail's main partner in crime, alongside Salleh. He appears bubbly and friendly when he is not disguised as the alleged Durian Ghost.
Tauke (a term used for Chinese business owners)	Male	Chinese	The businessman who purchases illegally captured animals from Pak Mail. He appears off-screen briefly through a phone call, but his voice indicates that he is a Chinese male.
Pak Cik Bas (bus driver)	Male	Malay	Drives the bus that Badrol and Lim take to Kampung Durian Runtuh. He has a very unbothered personality as he simply chills by the roadside when the bus breaks down.
Anak Pak Cik Bas (bus driver's son)	Male	Malay	Frustrated by his father, who treats him like a servant when the bus breaks down.
Mechanic	Male	Unknown	Appears off-screen as a mere voice on the phone when the bus driver calls him for repair enquiries.
Rokiah	Female	Malay	A news anchor who interviews Tok Dalang about missing durians in the village.
Crew	Unknown	Unknown	An off-screen camera crew member who silently nods as a response to Tok Dalang during his interview.

Driving Instructor	Male	Malay	Briefly appears during the motor race scene between Badrol and Lim. He is a licensed driving instructor who is in the middle of teaching his student how to drive when Badrol and Lim disrupt their lesson with their race.
Driving Student	Female	Malay	A scared student learning to drive.
Snake	Unknown	Animal (Snake)	One of the creatures in the movie that chases the characters and gets into a fight with Oopet's mother.

846

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- MovieStats.xlsx
- SeriesStatsSeason1andSeason2.xlsx