

The Study of Communication Social Networks in The Marvel Universe by Using a Social Network Analysis

การศึกษาเครือข่ายสังคมของตัวละครในหนังสือการ์ตูนมาร์เวล

ด้วยการวิเคราะห์เครือข่ายสังคม

ธนัช เบญจอนูอาชา

รหัสนักศึกษา 63199130111

ภาควิชาวิทยาการคอมพิวเตอร์,

คณะวิทยาศาสตร์

มหาวิทยาลัยศรีนครินทรวิโรฒ

ประสานมิตร

กรุงเทพฯ, ประเทศไทย

e-mail : thanus.ben@g.swu.ac.th

บทคัดย่อ--การวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis) การวิจัยนี้เป็นการวิจัยเชิงพรรณนา มีวัตถุประสงค์เพื่อศึกษาลักษณะเครือข่ายสังคมระหว่างตัวละครมาร์เวล และ ระหว่างหนังสือการ์ตูนและการปรากฏตัวในหนังสือการ์ตูนของตัวละครมาร์เวล วิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนาและการวิเคราะห์เครือข่ายสังคมด้วยข้อมูลเชิงสัมพันธ์และแผนภาพเครือข่าย

ผลการวิจัย พบว่าตัวละครที่มี Degree Centrality ปรากฏในหนังสือการ์ตูน(Comics) มากที่สุด 5 อันดับแรกได้แก่ SPIDER-MAN/PETER PARKER, CAPTAIN AMERICA, IRON MAN/TONY STARK, THING/BEJAMIN J. GR, THOR/DR. DONALD BLAK ตามลำดับ และหนังสือการ์ตูน (comics) ที่มี Degree Centrality ตัวละคร (Hero) ปรากฏตัวมากที่สุด 5 อันดับแรกได้แก่ Contest of Champions #1 (COC 1), Infinity War #3 (IW 3), Infinity War #1 (IW 1), H2 279, Infinity War #4 (IW 4) ตามลำดับ และตัวละครที่มีความสัมพันธ์ (Relations) กับตัวละครอื่นมากที่สุด 5 อันดับแรก CAPTAIN AMERICA, SPIDER-MAN/PETER, IRON MAN/TONY STARK, WOLVERINE/LOGAN, และ SCARLET WITCH/WANDA ตามลำดับ (Abstract)

Keywords—social network, marvel universe, Communication Social Networks, Hero , Hero marvel ,Comic (key words)

I. บทนำ

มาร์เวลคอมิกส์ (อังกฤษ: Marvel Comics) เป็นชื่อแบรนด์และตราประทับหลักของ บริษัท มาร์เวลเวิลด์ไวด์ (Marvel Worldwide Inc.) เดิมชื่อ บริษัท มาร์เวลพับลิชชิง (Marvel Publishing, Inc.) และ มาร์เวลคอมิกส์กรุ๊ป (Marvel Comics Group) ซึ่งเป็นผู้จัดพิมพ์หนังสือการ์ตูนอเมริกันและสื่อที่เกี่ยวข้อง ในปี ค.ศ. 2009 เดอะวอลต์ดิสนีย์ ได้เข้าซื้อกิจการมาร์เวลเอนเตอร์เทนเมนต์ ซึ่งเป็นบริษัทแม่ของ Marvel Worldwide

มาร์เวลคอมิกส์เป็นค่ายการ์ตูนและสื่อต่าง ๆ ที่เกี่ยวข้องกับซูเปอร์ฮีโรของสหรัฐอเมริกา ก่อตั้งในปี ค.ศ. 1939 โดย มาร์ติน กูดแมน ในนามของไทม์ลีคอมิกส์ มีนักเขียน นักวาดคนสำคัญ เช่น สแตน ลี, แจ็ก เคอร์บี้, สตีฟ ดิตโก เป็นต้น มาร์เวลคอมิกส์ มีชื่อเสียงโด่งดังและรู้จักกันดี เช่น เอ็กซ์เมน, สไปเดอร์-แมน, ฮัลก์, กัปตันอเมริกา, ไรออนแมน, ธอร์ เป็นต้น และศัตรูที่โด่งดังและรู้จักกันดี เช่น กรีนก๊อบลิน, แมกนีโต, ดอกเตอร์ดรัม, โลกิ, กาแล็กตัส, และเรดสกาย เป็นต้น มาร์เวลคอมิกส์มีคู่แข่งที่สำคัญมา

ยาวนานคือดิจิทัลคอมมิก ปัจจุบันมาร์เวลคอมิกส์เป็นส่วนหนึ่งของ บริษัทเดอะวอลต์ดิสนีย์

ในปัจจุบัน มีหนังสือการ์ตูน (Comic) ออกมามากมาย ใน dataset ที่ได้นำมาวิเคราะห์มีถึง 12,651 เล่ม และมีตัวละครกว่า 6,439 ตัวละคร จึงเป็นที่น่าสนใจในการนำ data set มาวิเคราะห์ถึง เครือข่ายสังคม (Social network) ซึ่ง data set จะประกอบไปด้วย เครือข่ายสังคม (Social network) รายชื่อตัวละคร (Hero) กับ รายชื่อ ตอนของหนังสือการ์ตูน (Comic) หมายถึงตัวละครที่ปรากฏตัวใน หนังสือการ์ตูนในตอนนั้นๆ และ เครือข่ายสังคม (Social network) ของ ตัวละคร (Hero) กับ ตัวละคร (Hero) หมายถึงความสัมพันธ์ ระหว่างตัวละคร (Relationship)

II. TECHNICAL BACKGROUND

โครงสร้างพื้นฐานหลักในการสร้างกราฟ ประกอบไปด้วย 4 ส่วนได้แก่

1. โหนด (Node) โดยปกติจะใช้เพื่อเก็บข้อมูลในฐานะข้อมูลของ กราฟ หากเปรียบกับระบบฐานข้อมูลเชิงสัมพันธ์ก็เหมือนกับเอนทิตี (Entity) ในระบบฐานข้อมูลเชิงสัมพันธ์นั่นเองในที่นี้มี โหนด ของตัวละคร Hero และ โหนดของหนังสือการ์ตูน

2. ความสัมพันธ์ (Relationship) เป็นการแสดงการเชื่อมต่อ โหนดกับอีกโหนดหนึ่งหากเปรียบกับระบบฐานข้อมูลเชิงสัมพันธ์ ก็เหมือนกับความสัมพันธ์ระหว่างตาราง (Table) หรือระหว่างเอนทิตี (Entity) แต่ในความสัมพันธ์ของกราฟจะอำนวยความสะดวกในการสอบถามข้อมูล(Query)ในลักษณะของสำรวจเส้นทาง (Traversal) ซึ่งแสดงความสัมพันธ์ของฐานข้อมูลกราฟจะใช้เพื่อ แสดงการเริ่มต้นและสิ้นสุดโหนดแสดงทิศทางหรือไม่มีทิศทาง ของหัวลูกศรของกราฟ หรือเรียกว่า Edge นอกจากนี้ความสัมพันธ์ ยังสามารถอ้างอิงความสัมพันธ์เองหรือวนลูปได้อีกด้วย

3. คุณสมบัติ (Property) ทั้งโหนดและความสัมพันธ์จะต้องมี คุณสมบัติเพื่อแสดงถึงข้อมูลของตนเองโดยคุณสมบัติจะเก็บข้อมูล แบบคู่ข้อมูลประกอบด้วยชื่อและค่าของข้อมูลคู่กัน(name/value pairs) หากเปรียบกับระบบฐานข้อมูลเชิงสัมพันธ์ก็เหมือนกับการ เก็บค่าของฟิลด์หรือแอตทริบิวต์

4.ป้ายกำกับหรือลาเบล (Label) โดยมีไว้เพื่อเพิ่มความรวดเร็ว ในการค้นหา และการสร้างกราฟได้อย่างมีประสิทธิภาพ สำหรับ Neo4j

รูปแบบของเครือข่าย (Network) แบ่งออกได้เป็น 2 รูปแบบ ดังนี้

1. **Undirected Network** คือระบบเครือข่ายที่ไม่มีทิศทางของการเชื่อมความสัมพันธ์ เช่น ข้อมูล A,B และ C Node ทั้ง 3 นี้มีความสัมพันธ์กันทั้งหมด

2. **Directed Network** คือระบบเครือข่ายที่มีทิศทางของการเชื่อมความสัมพันธ์กัน เช่น A มีความสัมพันธ์ชี้ไปหาทั้ง B และ C แต่ C ไม่ได้มีความสัมพันธ์ไปหาทั้ง A และ B

อัลกอริทึม (Algorithm) ที่ใช้ในการวิเคราะห์

ระยะทางที่สั้นที่สุด

Shortest Path คือ การค้นหาเส้นทางระหว่างจุดข้อมูล 2 จุด (Node) ในกราฟเพื่อลดผลรวมของน้ำหนักของขอบที่เป็นส่วนประกอบ ซึ่งเป็นการเดินทางโดยใช้ระยะทางที่สั้นที่สุดและใช้เวลาน้อยที่สุด ตัวอย่างเช่น การแนะนำเพื่อน (Friend Suggestion) บน Facebook หรือการวิเคราะห์เส้นทางการส่งของของบริษัทขนส่งมักจะใช้ หลักการนี้มาวิเคราะห์เพื่อประเมินการจัดส่งให้คุ้มค่าที่สุดและ ค่าใช้จ่าย เป็นต้น

การหาความเป็นจุดศูนย์กลาง(Centrality)

Centrality เป็นการวิเคราะห์หาความสำคัญของจุดข้อมูล (Node) ที่มีความสำคัญมากที่สุดภายในเครือข่าย (Network) ซึ่งกล่าวคือ สามารถนำมาใช้ในการระบุตัวบุคคลที่มีอิทธิพลมากที่สุด ใน เครือข่ายทางสังคมเป็นต้นโดยในบทความนี้จะใช้การวิเคราะห์จาก วิธี Degree Centrality, Betweenness Centrality และ Closeness Centrality โดยในแต่ละอัลกอริทึมสามารถอธิบายรายละเอียดในเบื้องต้นได้ดังนี้

1. **Degree Centrality** คือ เป็นการจัดอันดับของโหนดด้วยการ เชื่อมต่อที่มากกว่าในลักษณะของการเป็นศูนย์กลางของข้อมูล ซึ่ง เป็นอัลกอริทึมที่ช่วยให้ง่ายในการค้นหาโหนดที่สำคัญในเครือข่าย ของข้อมูล อีกทั้งช่วยในการวิเคราะห์เอนทิตีที่มีอิทธิพลโดยพิจารณา จากจำนวนความสัมพันธ์ที่โดยรอบ

2. **Betweenness Centrality** คือ เป็นการวิเคราะห์หาจำนวนของ โหนดที่มีอิทธิพลที่สุดโดยเป็นจุดศูนย์กลางของข้อมูลโดยพิจารณา จากความสำคัญของโหนดที่เชื่อมต่อโหนดอื่นๆ ซึ่งมักใช้เพื่อ ค้นหาโหนดที่ทำหน้าที่เป็นสะพานเชื่อมต่อกับส่วนหนึ่งของกลุ่ม ข้อมูลไปยังอีกกลุ่มหนึ่ง โดยเป็นอัลกอริทึมที่คำนวณเส้นทางที่สั้น ที่สุดระหว่างโหนดทุกคู่ที่มีการเชื่อมต่อกัน

3. **Closeness Centrality** คือ การพิจารณาข้อมูลของโหนด (Node) ที่อยู่จุดศูนย์กลางว่ามีความใกล้ชิดกับ Node ต่างๆที่อยู่ภายในกลุ่มเดียวกัน ถ้าโหนด (Node) นั้นเข้าเงื่อนไขดังกล่าวแสดงว่า Node นั้นสามารถส่งข้อมูลและสื่อสารกับ Node อื่นๆได้คล่องตัวคือสามารถเป็นตัวกระจายข้อมูลหรือข่าวสารได้อย่างมีประสิทธิภาพมากที่สุด

III. METHODS

ขั้นตอนกระบวนการนี้เราเอา data set มาจาก www.kaggle.com ข้อมูลมีทั้งสิ้น 3 ไฟล์ คือ edges.csv, hero-network.csv และ nodes.csv จะขออธิบายไฟล์แรกก่อน edges.csv ซึ่งจะเป็นข้อมูลความสัมพันธ์ตัวละคร (hero) กับหนังสือการ์ตูน (comic) จะบอกความสัมพันธ์ของตัวละครที่ปรากฏในหนังสือการ์ตูนเล่มที่เท่าไร ไฟล์ที่ 2 hero-network.csv จะเป็นข้อมูลความสัมพันธ์ระหว่างตัวละครกับตัวละคร และไฟล์สุดท้าย edges.csv จะเป็นเหมือน data dictionary

ในการวิเคราะห์เราจะใช้โปรแกรม neo4j เป็นหนึ่งในฐานข้อมูลที่กำลังนิยมในปัจจุบัน เพราะความยืดหยุ่นในการใช้งานนั้นค่อนข้างดีกว่า NoSql ทั่วๆไป

Neo4J เป็นฐานข้อมูลประเภท Graph Database(GDBMS) ถูกพัฒนามาบนพื้นฐานภาษา JAVA และใช้ภาษา CQL(Cypher Query Language Neo4j) ในการ Query ข้อมูลต่างๆ

โดยได้เอาข้อมูล edges.csv และ hero-network.csv import เข้าไปใน neo4j และทำการวิเคราะห์และสร้างกราฟต่างๆ

IV. EXPERIMENTAL RESULTS

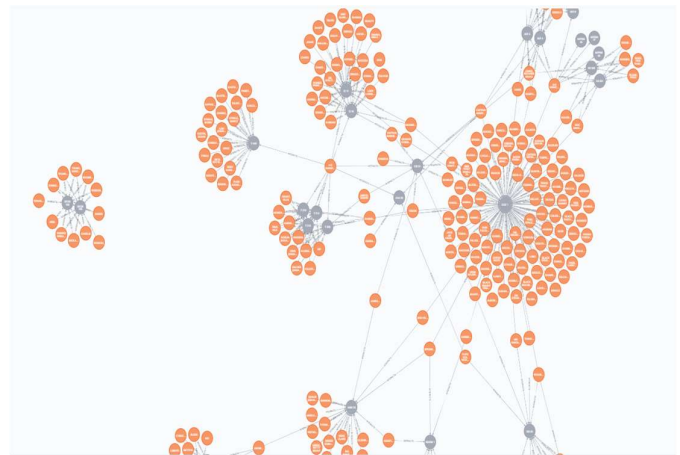
จากการวิเคราะห์ข้อมูล จะประกอบไปด้วยโหนด (Node) 4 โหนด ดังนี้ โหนด Hero และโหนด Comic ซึ่ง 2 โหนดนี้เป็นความสัมพันธ์กันที่มากจากไฟล์ edges.csv และ โหนด hero1 และ โหนด hero2 ที่มีความสัมพันธ์กันระหว่างตัวละครที่มาจากไฟล์ hero-network.csv ซึ่งผลที่ได้จากการวิเคราะห์ได้ดังนี้



<



โหนด Hero จะมีข้อมูลตัวละคร 6,439 โหนดหมายถึงมีตัวละคร 6,439 ตัวละคร, โหนด comic มีข้อมูลหนังสือการ์ตูน 12,651 โหนด, โหนด Hero1 มีตัวละคร 6,211 โหนด, Hero2 มีตัวละคร 6,173 โหนด



ในแต่ละประเภทโหนด ได้ค่าสถิติเบื้องต้นดังนี้

โหนด hero

	min	max	avg_comic	stdev
1	1	1577	14.925298959465763	59.26872510194238

จากรูปโหนด Hero มีค่า min เป็น 1, Max 1,577, ค่าเฉลี่ย 14.93, ค่าเบี่ยงเบน 59.27

อธิบายคือตัวละครที่ปรากฏตัวในหนังสือการ์ตูนน้อยสุด (min) คือ 1 เล่ม และปรากฏตัวมากที่สุด (max) คือ 1,577 เล่ม ค่าเฉลี่ยที่ตัวละครทั้งหมดปรากฏตัวในหนังสือการ์ตูนคือ 14.93 เล่ม จะเห็นได้ว่าค่าเฉลี่ยค่อนข้างต่างค่า max ค่อนข้างมาก ซึ่งก็เป็นผลให้ค่าเบี่ยงเบนค่อนข้างเยอะตาม

โหนด comic

	min	max	avg	stdev
1	1	111	7.596553632123889	6.438439465127829

จากรูปโหนด comic มีค่า min เป็น 1, Max 111, ค่าเฉลี่ย 7.60, ค่าเบี่ยงเบน 6.44

อธิบายคือการ์ตูนที่มีตัวละครที่ปรากฏตัวน้อยสุด (min) คือ 1 ตัวละคร และมากที่สุด (max) คือ 111 ตัวละคร ค่าเฉลี่ยหนังสือการ์ตูนเล่มหนึ่งจะมีตัวละครเฉลี่ยที่ 7.60 ตัวละคร

โหนด hero1

	min	max	avg_hero	stdev
1	1	1426	36.094187731444514	83.23312419338517

จากรูปโหนด hero1 มีค่า min เป็น 1, Max 1,426, ค่าเฉลี่ย 36.09, ค่าเบี่ยงเบน 83.23

อธิบายคือตัวละครที่มีความสัมพันธ์กับตัวละครอื่นๆ น้อยสุด (min) คือ 1 ตัวละคร และมากที่สุด (max) คือ 1,426 ตัวละคร ค่าเฉลี่ยความสัมพันธ์ของตัวละครเฉลี่ยที่ 36.09 ตัวละคร

โหนด hero2

	min	max	avg_hero	stdev
1	1	1428	36.31637777417785	83.1387906362528

จากรูปโหนด hero2 มีค่า min เป็น 1, Max 1,428, ค่าเฉลี่ย 36.32, ค่าเบี่ยงเบน 83.14

และในทำนองเดียวกันตัวละครที่มีความสัมพันธ์กับตัวละครอื่นๆ น้อยสุด (min) คือ 1 ตัวละคร และมากที่สุด (max) คือ 1,428 ตัวละคร ค่าเฉลี่ยความสัมพันธ์ของตัวละครเฉลี่ยที่ 36.32 ตัวละคร

Centrality Degree

Centrality degree ของ hero

hero	degree
1 "SPIDER-MAN/PETER PARKER"	1577
2 "CAPTAIN AMERICA"	1334
3 "IRON MAN/TONY STARK"	1150
4 "THING/BENJAMIN J. GR"	963
5 "THOR/DR. DONALD BLAK"	956

จากตาราง Centrality degree ของ hero 5 อันดับแรกมีจำนวน ดังนี้

1.ตัวละคร "SPIDER-MAN/PETER PARKER" มีจำนวน degree 1,577, 2. "CAPTAIN AMERICA" มีจำนวน degree 1,334, 3. "IRON MAN/TONY STARK" มีจำนวน degree 1,150, 4. "THING/BENJAMIN J. GR" มีจำนวน degree 963, 5. "THOR/DR. DONALD BLAK" มีจำนวน degree 956

จะเห็นได้ว่า SPIDER-MAN/PETER PARKER เป็นตัวละครที่สำคัญที่สุด จำนวน degree 1,577 หมายถึง ตัวละคร SPIDER-MAN/PETER PARKER ได้ปรากฏตัวในหนังสือการ์ตูน comic ถึง 1,577 เล่ม ตามมาด้วยตัวละคร "CAPTAIN AMERICA" "IRON MAN/TONY STARK" "THOR/DR. DONALD BLAK" "THING/BENJAMIN J. GR" ตามลำดับ จะสังเกตได้ว่า



ตัวละครส่วนใหญ่เป็นตัวละครที่ถูกนำไปสร้างเป็นหนัง มีเพียง "THING/BENJAMIN J. GR" เท่านั้นที่ยังไม่ถูกนำไปสร้างเป็นหนัง

จากรูปด้านบน คือตัวละคร "THING/BENJAMIN J. GR"

Centrality degree ของ comic

comic	degree
1 "COC 1"	111
2 "IW 3"	91
3 "IW 1"	90
4 "H2 279"	87
5 "IW 4"	80

จากตาราง Centrality degree ของ comic 5 อันดับแรกมีจำนวน ดังนี้

1.หนังสือการ์ตูน "COC 1" มีจำนวน degree 111, 2. "IW 3" มีจำนวน degree 91, 3. "IW 1" มีจำนวน degree 90, 4. "H2 279" มีจำนวน degree 87, 5. "IW 4" มีจำนวน degree 80

หนังสือการ์ตูน "COC 1" มีจำนวน degree 111 ซึ่งหมายถึง มีตัวละครที่ปรากฏตัวในหนังสือการ์ตูนเล่มนี้ถึง 111 ตัว ตามมาด้วย "IW 3", "IW 1", "H2 279", "IW 4" ตามลำดับ



รูป หนังสือการ์ตูน "COC 1"

Centrality degree ของ hero1

	hero1	degree
1	"CAPTAIN AMERICA"	1426
2	"SPIDER-MAN/PETER PAR"	1286
3	"IRON MAN/TONY STARK "	1132
4	"WOLVERINE/LOGAN "	1036
5	"SCARLET WITCH/WANDA "	1004

จากตาราง Centrality degree ของข้อมูล hero1 5 อันดับแรกมีจำนวน ดังนี้

1.ตัวละคร "CAPTAIN AMERICA" มีจำนวน degree 1,426, 2. "SPIDER-MAN/PETER PARKER" มีจำนวน degree 1,286, 3. "IRON MAN/TONY STARK" มีจำนวน degree 1,132, 4."WOLVERINE/LOGAN " มีจำนวน degree 1036, 5. "SCARLET WITCH/WANDA " มีจำนวน degree 1,004

ในส่วนนี้จะหมายถึงตัวละครที่มีความสัมพันธ์กับตัวละครอื่นๆ "CAPTAIN AMERICA" มีความสัมพันธ์กับตัวละครอื่นๆถึง 1,426 ตัวละคร และตามมาด้วย "SPIDER-MAN/PETER PARKER", "IRON MAN/TONY STARK", ". "WOLVERINE/LOGAN", "SCARLET WITCH/WANDA" ตามลำดับ

Centrality degree ของ hero2

	hero2	degree
1	"CAPTAIN AMERICA"	1428
2	"SPIDER-MAN/PETER PAR"	1277
3	"IRON MAN/TONY STARK "	1106
4	"THING/BENJAMIN J. GR"	1020
5	"MR. FANTASTIC/REED R"	1001

จากตาราง Centrality degree ของข้อมูล hero2 5 อันดับแรกมีจำนวน ดังนี้

1.ตัวละคร "CAPTAIN AMERICA" มีจำนวน degree 1,428, 2. "SPIDER-MAN/PETER PARKER" มีจำนวน degree 1,277, 3. "IRON MAN/TONY STARK" มีจำนวน degree 1,106, 4. "THING/BENJAMIN J. GR" มีจำนวน degree 1020, 5. "MR. FANTASTIC/REED R" มีจำนวน degree 1,001

V. CONCLUSIONS

การวิเคราะห์เครือข่ายสังคมของตัวละครในหนังสือการ์ตูนมาร์เวลนี้เป็นการรื้อข้อมูล โดยเป็นข้อมูล data source public ที่เอามาจาก www.kaggle.com และได้นำมาวิเคราะห์กราฟผ่านโปรแกรม Neo4j ฐานข้อมูลประเภท Graph Database(GDBMS) ถูกพัฒนามาบนพื้นฐานภาษา JAVA และใช้ภาษา CQL(Cypher Query Language Neo4j) ในการ Query ข้อมูลต่างๆ

การปรากฏตัวของตัวละครกับหนังสือการ์ตูน comic และความสัมพันธ์ระหว่างตัวละคร วัดจำนวนความสัมพันธ์จากค่า Centrality degree โดยเราได้รู้ว่าจากข้อมูลที่นำมาวิเคราะห์ทั้งหมด ตัวละคร "SPIDER-MAN/PETER PARKER" มีความสำคัญมีการปรากฏตัวในหนังสือการ์ตูน (Comic) มากที่สุดเป็นลำดับแรก และได้ทราบว่าหนังสือการ์ตูน "COC 1" เป็นหนังสือการ์ตูนที่มีตัวละครปรากฏเยอะที่สุด และในเรื่องของความสัมพันธ์ระหว่างตัวละคร ได้พบว่าตัวละคร "CAPTAIN AMERICA" มีความสัมพันธ์เกี่ยวข้องกับตัวละครอื่นๆมากที่สุด

VI. REFERENCES

รศ.ดร. วุฒิชัย ร่มสายหยุด, นิโอโฟร์เจ ฐานข้อมูลกราฟสำหรับการวิเคราะห์ - กรุงเทพฯ : ซีเอ็ดดูเกชั่น, 2563
[HTTPS://WWW.KAGGLE.COM/CSANHUEZA/THE-MARVEL-UNIVERSE-SOCIAL-NETWORK](https://www.kaggle.com/CSANHUEZA/THE-MARVEL-UNIVERSE-SOCIAL-NETWORK)
[HTTPS://TH.WIKIPEDIA.ORG/WIKI/%E0%B8%A1%E0%B8%B2%E0%B8%A3%E0%B9%8C%E0%B9%80%E0%B8%A7%E0%B8%A5%E0%B8%84%E0%B8%AD%E0%B8%A1%E0%B8%B4%E0%B8%81%E0%B8%AA%E0%B9%8C](https://th.wikipedia.org/wiki/%E0%B8%A1%E0%B8%B2%E0%B8%A3%E0%B9%8C%E0%B9%80%E0%B8%A7%E0%B8%A5%E0%B8%84%E0%B8%AD%E0%B8%A1%E0%B8%B4%E0%B8%81%E0%B8%AA%E0%B9%8C)