فصل دهم

موضوع : Mining Social-Network Graphs

به نام خدا

مقدار زیادی داده با تجزیه و تحلیل از شبکه های اجتماعی به دست می آید.

شناخته شده ترین نمونه یک شبکه اجتماعی ، رابطه "دوستان" است که در سایت هایی مانند فیس بوک یافت می شود.

با این حال ، همانطور که خواهیم دید بسیاری از منابع داده دیگر وجود دارند که افراد یا نهادهای دیگر را به هم وصل می کنند.

در این فصل تکنیک های تحلیل داده های موجود در شبکه ها را بررسی می کنیم.

یک سوال مهم در مورد یک شبکه اجتماعی ، چگونگی شناسایی "اجتماعات" ، یعنی زیر مجموعه گره ها (افراد یا اشخاص دیگری که شبکه را تشکیل می دهند) با اتصالات غیرمعمول قوی است.

برخی از تکنیک های مورد استفاده برای شناسایی جوامع مشابه الگوریتم های خوشه بندی است که در فصل 7 درباره آنها گفتیم.

با این حال ، مجموعه ها تقریباً هرگز قسمت هایی از گره ها را در یک شبكه تقسیم نمی كنند.

در عوض ، مجموعه ها معمولاً با هم همپوشانی دارند.

به عنوان مثال ، شما ممکن است به چندین گروه مختلف از دوستان یا همکلاسی های خود تعلق داشته باشید.

مردم یک جامعه تمایل دارند یکدیگر را بشناسند ، اما مردم دو جامعه مختلف به ندرت یکدیگر را می شناسند.

شما نمی خواهید فقط در یکی از گروه ها قرار بگیرید و همچنین معقول نیست که همه ی افراد از همه ی جوامع خود را حتما در یک گروه قرار دهید.

همچنین در این فصل الگوریتم های کارآمد را برای کشف سایر خصوصیات گراف ها بررسی می کنیم.

ما به یک مفهومی تحت عنوان simrank نگاه می کنیم که این مفهوم به معنی کشف شباهت بین گره های یک گراف است.

ما شمارش مثلثی را به عنوان روشی برای سنجش ارتباط در یک جامعه استفاده می کنیم.

ما برای اندازه گیری دقیق و تقریبی اندازه مجموعه ها و گره های در یک گراف ، الگوریتم های مناسبی را ارائه می دهیم.

سرانجام ، برای محاسبه بستار انتقالی ، به الگوریتم های ویژه ای را بررسی می کنیم.