
بسم الله الرحمن الرحیم

دانشگاه تهران

پردیس علوم

دانشکدهٔ ریاضی، آمار و علوم کامپیوتر

عنوان

اندیس وینر درختان دودویی متعادل

نگارش

مهدی روستا

استاد راهنما

دکتر حسن یوسفی آذری

پایان‌نامه

جهت دریافت درجه کارشناسی ارشد

در ریاضی کاربردی

تیر ۸۷

تقدیم به پدر و مادر عزیزم

که در راه علم و دانش از هیچ کمکی به اینجانب دریغ نکردند،

تقدیم به همسر مهربانم

که با صبر و فداکاری همواره مشوق پیشرفت اینجانب بوده است،

تقدیم به برادر و خواهر عزیزم

دکتر محسن روستا که با راهنماییشان به بنده کمک کردند و خواهرم ملیکا روستا که با دست‌های کوچکش و قلب بزرگش برایم دعا می‌کرد.

تقدیر و تشکر

از تمام کسانی که برای علم احترام قائلند و تمام کسانی که به اینجانب کمک کردند اعم از خانواده‌ها، دوستان، اطرافیان، معلمان و استادانم که در تمام دوران تحصیلم باعث پیشرفت اینجانب شده‌اند به خصوص معلم گرامی آقای نظری و استاد ارجمندم دکتر حسن یوسفی آذری کمال تشکر را دارم.

چکیده

اندیس‌های توپولوژیک برای گراف‌ها چندی است که مورد بررسی و علاقه محققان قرار گرفته است و تاکنون اندیس‌های مختلفی برای گراف‌های مولکولی ترکیبات شیمیایی ارائه شده است. اندیس وینر و میانگین فاصله برای دیگر علوم پرکاربرد هستند. در واقع، این هارولد وینر بود که در سال ۱۹۴۷ معین کرد که W یک اندازه برای درجه شاخه‌های مولکولی است که به نظر می‌رسید وابسته به خیلی از خواص فیزیکی و شیمیایی مولکول باشد. تعداد زیادی مقاله در مورد اندیس وینر و میانگین فاصله هم در مقالات ریاضی و هم در مقالات شیمی چاپ شده است [7,12]. در این پایان نامه چندین مقاله را که اخیرا در باره اندیس وینر درختان در مجلات معتبر انتشار یافته است مورد بررسی و تجزیه و تحلیل قرار می‌دهیم.

پیش‌گفتار

ریاضیات به تعبیری مادر علوم و نظریه گراف یکی از پرکاربردترین شاخه‌های این علم می باشد. این شاخه، که تا حدود نیم قرن پیش، مهجور بود امروزه کاربردهای گسترده‌ای در درون و برون ریاضیات دارد. به تعبیری می‌توان گفت یکی از قوی‌ترین ابزارهای ارتباط دهنده بین شاخه‌های علوم می‌باشد. البته این تاثیرات دو سویه بوده است.

پایان نامه حاضر در چهار فصل تنظیم گردیده است، که فصل اول به بیان مقدمات و تعاریف اولیه از نظریه گراف می‌پردازد.

در فصل دوم اندیس وینر را تعریف و فرمول‌هایی بازگشتی و غیر بازگشتی برای درختان به دست می‌آوریم.

در فصل سوم اندیس وینر درختان n رأسی با ماکزیمم درجه را محاسبه می‌کنیم. در این فصل روی دو مسئله مهم کار می‌کنیم، اول اینکه چه درختانی با مرتبه n و ماکزیمم درجه از حداکثر Δ، اندیس وینر را می‌نیمم می‌کنند؟ و دوم اینکه چه درختانی از میان تمام درختان با مرتبه n، اندیس وینر را ماکزیمم می‌کنند که در آن رئوس یا یک رأس تنها یا از ماکزیمم درجه Δ هستند؟

در فصل چهارم کلاس جدیدی از درختان (kـدرختان) را معرفی کرده اعمال خاصی را روی آنها انجام می‌دهیم و سپس اندیس وینر درختان تغییر یافته را به دست می‌آوریم و حدس‌هایی مهم می‌زنیم و برای اثبات یا رد آنها از الگوریتم‌ها و آزمایشاتی استفاده می‌کنیم و در نهایت اندیس وینر درختان دودویی متعادل را محاسبه می‌کنیم.

فهرست مطالب

فصل اول

تعاریف و نمادگذاری

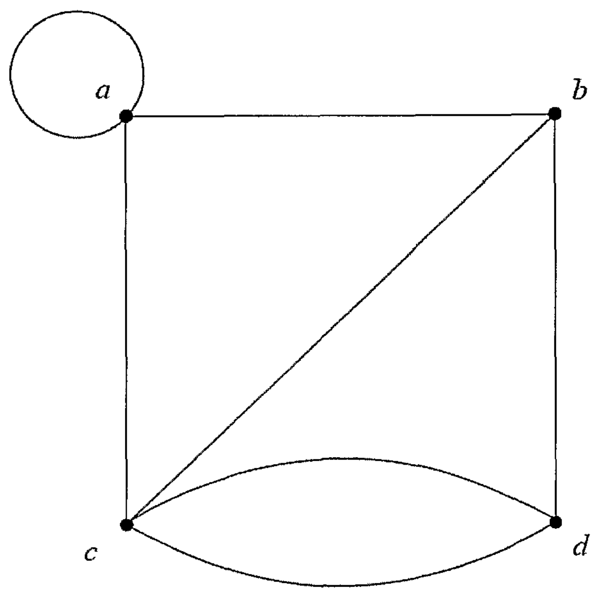
۱-۱ مقدمه

در دنیای اطراف ما، اوضاع فراوانی وجود دارند که می‌توان توسط نموداری متشکل از یک مجموعه نقاط، به علاوه خطوطی که برخی از این نقاط را به یکدیگر متصل می‌کنند به توصیف و بررسی آنها پرداخت. به عنوان مثال برای نشان دادن یک شبکه ارتباطی می‌توان از نموداری استفاده کرد که در آن، نقاط نمایانگر مراکز ارتباطی و خطوط نشان دهنده پیوندهای ارتباطی بین مراکز می‌باشند. توجه داشته باشید که در این گونه نمودارها، آنچه بیشتر مورد توجه است این است که آیا دو نقطه داده شده به وسیله یک خط به یکدیگر

متصل هستند یا نه؟ تعبیر ریاضی این اوضاع به مفهوم گراف منتهی می‌شود.

۱-۲ تعاریف مقدماتی گراف

یک چند گراف (گراف چندگانه) عبارت است از یک زوج مانند $G = (V(G), E(G))$ که در آن $V(G)$ یک مجموعه ناتهی از عناصری به نام رئوس یا نقاط است و $E(G)$ گردایه‌ای از زوج‌های نامرتب $V(G)$ (که لزوما متمایز نیستند) به نام یال‌ها می‌باشند. $V(G)$ را مجموعه رئوس G نامیده و $n(G) = |V(G)|$ را مرتبه G می‌گوییم. $E(G)$ را مجموعه یال‌های G نامیده و $|E(G)|$ را اندازه G می‌گوییم. رأس u همسایه (یا هم‌جوار) رأس v می‌گوییم هرگاه $uv \in E(G)$ باشد. شکل ۱ـ۱ مثالی از یک چندگراف است که در آن $V(G) = \{a, b, c, d\}$ و $E(G) = \{aa, ab, ac, bc, bd, cd, cd\}$.



شکل ۱ـ۱.

چندگراف $G = (V(G), E(G))$ را گراف ساده می‌گوییم هرگاه عناصر $E(G)$ متمایز باشند و هر عضو $E(G)$ نیز از عناصر متمایز $V(G)$ تشکیل شده باشد.

در سراسر این پایان نامه گراف‌های ساده مورد نظر می‌باشند و آنها را به طور خلاصه گراف می‌نامیم.

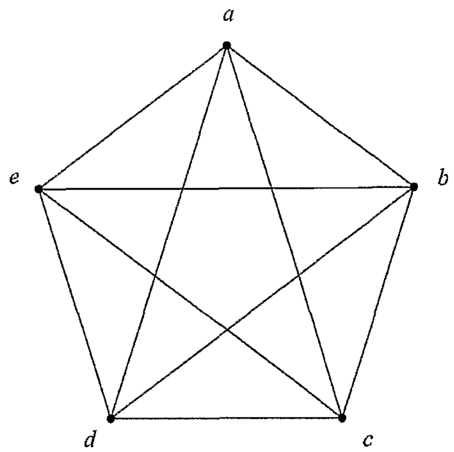
همسایگان رأس $v \in V(G)$ را با $N_G(v)$ نشان می‌دهیم. تعداد یال‌های گذرنده بر یک رأس را درجه آن رأس می‌نامیم. درجه رأس $v \in V(G)$ را با $d_G(v)$ نشان می‌دهیم و برابر $|N_G(v)|$ می‌باشد. یک رأس با درجه یک را، یک رأس تنها(برگ) می‌نامیم. ماکزیمم درجه رئوس گراف G را، $\max_{v\in V(G)} \deg_G(v)$ را با $\Delta(G)$ نشان می‌دهیم.

زیرگراف یک گراف مانند $G = (V(G), E(G))$ یک گراف مانند $G_1 = (V(G_1), E(G_1))$ است به طوری که $V(G_1) \subseteq V(G)$ و $E(G_1) \subseteq E(G)$. یک زیرگراف تولید شده (القایی) توسط مجموعه $X \subseteq V(G)$ از گراف G را با $G[X]$ نشان می‌دهیم.

دو گراف $G_1 = (V(G_1), E(G_1))$ و $G_2 = (V(G_2), E(G_2))$ را یکریخت می‌نامیم هرگاه تناظری یک به یک مانند σ از $V(G_1)$ به $V(G_2)$ (یا برعکس) وجود داشته باشد ($\sigma : V(G_1) \to V(G_2)$) به طوری که این تناظر مجموعه یال‌های $E(G_1)$ را به $E(G_2)$ تبدیل کند ($\sigma(E(G_1)) = E(G_2)$).

۱-۳ گشت، گذر، مسیر، دور یا مدار در گراف

بنا به تعریف یک گشت در چندگراف G عبارت است از دنباله‌ای از یال‌های G به طوری که رأس ابتدایی هر یال رأس انتهایی یال ماقبل باشد. بدیهی است در یک گشت یال‌ها لزوما متمایز نیستند. یک گشت را بسته می‌نامیم هرگاه رأس ابتدایی و انتهایی گشت یکی باشد. گشتی را که یال‌های آن متمایز باشند یک گذر می‌نامیم. بدیهی است در یک گذر رأس‌های لزوما متمایز نیستند. گشتی که تمام رئوس آن به جز رأس ابتدایی یا انتهایی گشت که می‌تواند یکی باشد متمایز باشند یک مسیر نام دارد. بدیهی است در هر مسیر یال‌ها نیز متمایز هستند. مسیر بسته را دور یا مدار می‌گوییم. تعداد یال‌های یک گشت را طول گشت می گوییم.



شکل ۱ـ۲.

در شکل ۱ـ۲، aedbeabde یک گشت و abcabd یک گذر است. همچنین edca یک مسیر و abcde یک مدار می‌باشد.

۱-۴ گراف‌های همبند

گراف G را همبند می‌گوییم هرگاه بین هر دو رأس دلخواه آن حداقل یک مسیر وجود داشته باشد، در غیر این صورت گراف را ناهمبند می‌گوییم. هر گراف ناهمبند از تعدادی گراف همبند تشکیل می‌شود که هر یک از آنها را یک مولفه از گراف می‌گوییم.

فاصله بین دو رأس u و v طول کوتاه‌ترین مسیر بین دو رأس u و v می‌باشد که با $d_G(u,v)$ نمایش داده می‌شود.

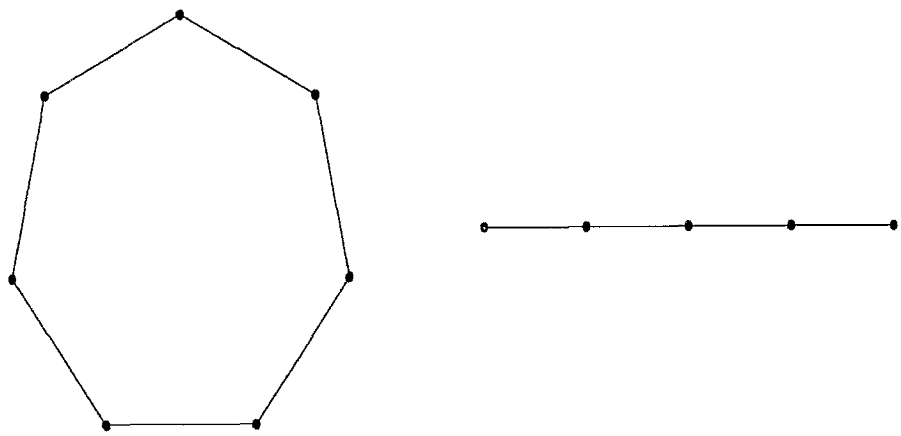
گراف همبند G را مداری گوییم هرگاه هر رأس آن دارای درجه دو باشد.

گراف مداری n رأسی را با C_n نشان می‌دهیم. بدیهی است که C_n دارای n یال است. زیرگراف حاصل از C_n پس از حذف یک یال گراف مسیری با n رأس نام دارد و با P_n نمایش داده می‌شود. P_n دارای $n-1$ یال است.

گراف همبند، فاقد مدار را درخت می‌نامیم. یک درخت n رأسی را با T_n نشان می‌دهیم. مجموعه رئوس T

را با $V(T)$ و مجموعه یال‌های آن را با $E(T)$ نشان می‌دهیم و $|V(T)| = n(T)$ و $|E(T)| = n(T) - ۱$ است.

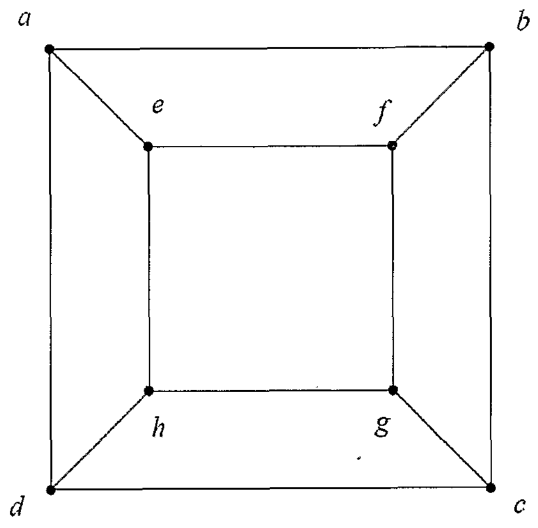
هر جفت رئوس از یک درخت با یک مسیر یکتا به هم متصل هستند. بدیهی است P_n یک درخت n رأسی است. درخت n رأسی (یکتا) با $n - ۱$ رأس تنها را ستاره می‌گوییم و با S_n نشان می‌دهیم.



شکل ۱ـ۳. $P_۵$ (سمت راست) و $C_۷$ (سمت چپ).

۱-۵ گراف دوبخشی

گراف $G = (V(G), E(G))$ را دوبخشی می‌نامیم هرگاه بتوان $V(G)$ را به دو مجموعه مانند $V_۱$ و $V_۲$ افراز نمود به طوری که هر یال از G یک سر در $V_۱$ و یک سر در $V_۲$ داشته باشد. به عنوان مثال گراف شکل ۱ـ۴ یک گراف دوبخشی است و در آن داریم:



$$V_1 = \{a, c, f, h\} \quad , \quad V_2 = \{b, d, e, g\}$$

شکل ۱ـ۴.

فصل دوم

اندیس وینر درختان

۲-۱ معرفی

فرض کنیم گراف G داده شده و G همبند باشد. فاصله بین رئوس u و v از G را با $d_G(u,v)$ نشان می‌دهیم. اندیس وینر G را با $W(G)$ نشان می‌دهیم و آن را به صورت زیر تعریف می‌کنیم:

$$W(G) = \sum_{\{u,v\} \subseteq V(G)} d_G(u,v) \qquad (۲\text{-}۱)$$

در نتیجه، میانگین فاصله بین رئوس G را با $\overline{d}(G)$[1] نشان می‌دهیم و برابر $\overline{d}(G) = \frac{W(G)}{\binom{n(G)}{2}}$ است.

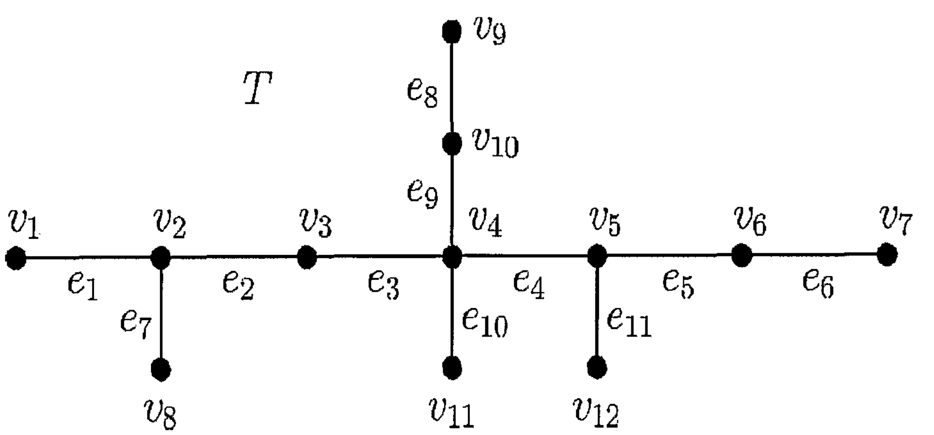
فاصله یک رأس $v \in V(G)$ از گراف G برابر جمع تمام فاصله‌ها بین رأس v و سایر رئوس G است و آن را با $d_G(v)$ نشان می‌دهیم.

اندیس وینر[2] یا عدد وینر برای اولین بار توسط هارولد وینر در سال ۱۹۴۷ معرفی شد. تعریف اندیس وینر به صورت جمع فاصله‌ها بین رئوس یک گراف به صورت معادله ۲-۱ اولین بار توسط هوسیا[3] [19] ارایه شد. درختان n رأسی (یکتا) با ۲ و $n-1$ رأس تنها را به ترتیب مسیر و ستاره می‌گوییم و با P_n و S_n نشان می‌دهیم. یک رأس v از درخت T را نقطه شاخه‌ای می‌گوییم هرگاه $d(v) \geq 3$ باشد. مسیرها، درختان بدون نقطه شاخه‌ای هستند.

دو نوع مسئله نزدیک به هم وجود دارد که توجه محققان را برای مدت طولانی به خود جذب کرده است:

چگونه $W(T)$ به ساختار T وابسته است؟

چگونه می‌توان به طور کارآیی، بدون استفاده از کامپیوتر، $W(T)$ را حساب کرد؟



شکل ۲ـ۱.

1) $\mu(G)$ 2) Wiener index or Wiener number 3) Hosoya

۲-۲ اندیس وینر برخی از درختان

برای تعداد زیادی از درختان (یا به طور دقیق‌تر کلاس‌های متفاوت از درختان) می‌توانیم اندیس وینر را به دست آوریم. شناخته‌ترین آنها

$$W(P_n) = \binom{n+۱}{۳} \quad , \quad W(S_n) = (n-۱)^۲ \tag{۲-۲}$$

هستند.

این فرمول‌ها در مقالات فراوانی آورده شده است. اگر T یک درخت n رأسی غیر از P_n و S_n باشد، آنگاه:

$$W(S_n) < W(T) < W(P_n) \tag{۳-۲}$$

در ادامه همین فصل اثبات آن را بیان خواهیم کرد.

فرض کنیم T_n با افزون یک رأس تنها به $T_{n-۱}$ به دست آمده باشد. از معادلات ۲-۲ و ۲-۳ نتیجه می‌شود که $W(T_n)$ حداقل به اندازه یک چندجمله‌ای درجه دو و حداکثر به اندازه یک چندجمله‌ای درجه سه از n افزایش می‌یابد.

دو خانواده از درختان دندریمر منتظم $T_{k,d}$ در شکل ۲ـ۲ نشان داده شده است که در آن d درجه رئوس غیر تنها و k فاصله بین رأس مرکزی و رئوس تنها را نشان می‌دهد. برای هر $d \geq ۳$، درخت $T_{k,d}$ دارای مرتبه

$$n(T_{k,d}) = ۱ + \frac{d}{d-۲}[(d-۱)^k - ۱]$$

است و اندیس وینر آن برابر زیر است:

$$W(T_{k,d}) = \frac{۱}{(d-۲)^۳}[(d-۱)^{۲k}[kd^۳ - ۲(k+۱)d^۲ + d] + ۲d^۲(d-۱)^k - d] \tag{۴-۲}$$

برای اعداد بزرگ k، $W \sim kn^۲$ است که به d $(d \geq ۳)$ وابسته نیست.

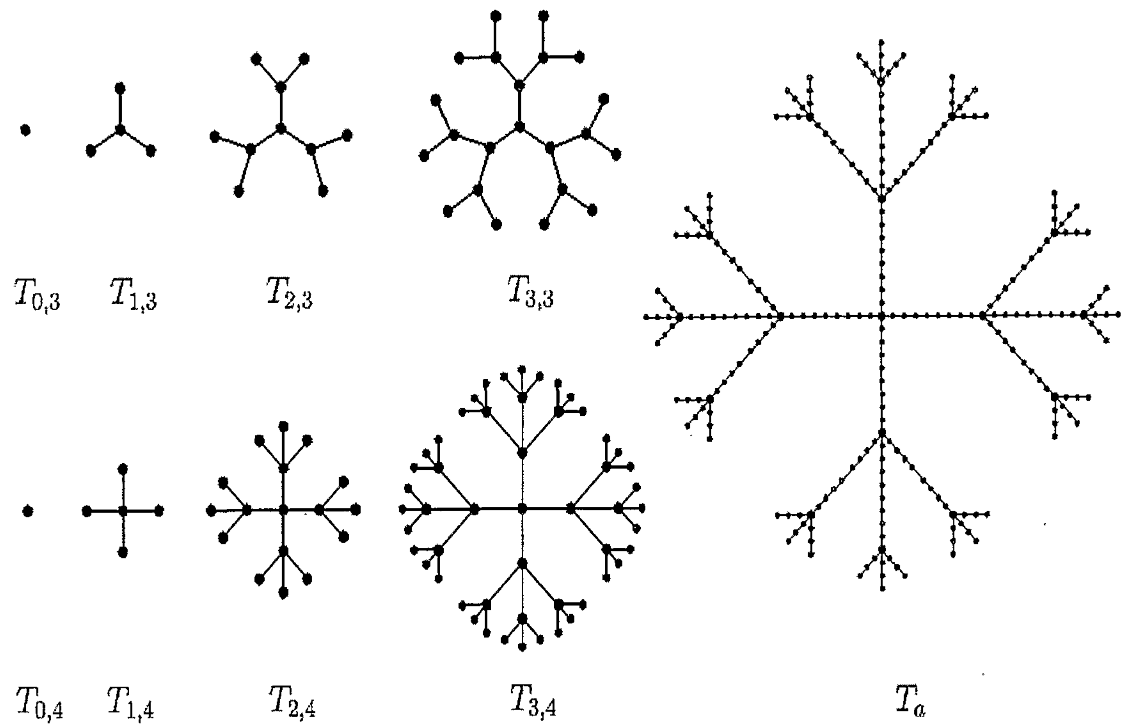
بیشترین کاربرد شیمیایی مورد استفاده برای $d = ۳, ۴$ می‌باشد که

$$W(T_{k,۳}) = (۹k - ۱۵)۲^{۲k} + ۱۸ \times ۲^k - ۳$$

و

$$W(T_{k,۴}) = (۴k - \frac{۷}{۲})۳^{۲k} + ۴ \times ۳^{k} - \frac{۱}{۲}$$

می‌باشد.



شکل ۲ـ۲.

۲-۳ مرکز و مرکزیت یک درخت

قبل از اینکه بحثمان را درباره تکنیک‌ها و فرمول‌ها، برای محاسبه اندیس وینر درختان ادامه دهیم، چند تعریف و خواص رئوس که نقش موثری در اندیس وینر دارد یادآوری می‌کنیم.

یک زیر درخت ماکزیمال شامل رأس v از درخت T به عنوان یک رأس تنها، یک شاخه از درخت T در رأس v نامیده می شود.

به عنوان مثال درخت T در شکل ۲ـ۱ دارای ۴ شاخه در رأس $v_۴$ می‌باشد. وزن شاخه B با $BW_T(B)$ نشان داده می‌شود و برابر تعداد یال‌های متعلق به آن می‌باشد. وزن شاخه‌ای رأس v با $BW_T(v)$ نشان داده می‌شود و برابر ماکزیمم وزن شاخه‌های متصل به v می‌باشد. بنابراین، برای شکل ۲ـ۱، $BW_T(v_۴) = ۴$ است. مرکزیت (شبه مرکز) درخت T را با $C(T)$ نشان می‌دهیم و برابر است با مجموعه رئوس T که دارای وزن شاخه‌ای می‌نیمم هستند. به عنوان مثال برای درخت T در شکل ۲ـ۱، $C(T) = \{v_۴\}$ است. جردن[4] مرکزیت یک درخت را معرفی کرد.

قضیه ۲-۱ اگر $C = C(T)$ مرکزیت درخت T از مرتبه n را نشان دهد، آنگاه یکی از موارد زیر صحیح است:

الف) $C = \{c\}$ و $BW(c) \leq \frac{n-۱}{۲}$.

ب) $C = \{c_۱, c_۲\}$ و $BW(c_۱) = BW(c_۲) = \frac{n}{۲}$.

و در هر مورد اگر $v \in V(T) \setminus C$ آنگاه $BW(v) \geq \frac{n}{۲}$ می‌باشد (اثبات در فصل ۳).

زلینکا[5] مجموعه رئوس با می‌نیمم فاصله را در درختان مشخص کرد [29].

قضیه ۲-۲ مجموعه رئوس با می‌نیمم فاصله در یک درخت T، مرکزیت T است.

نتیجه ۲-۱ رأس v از درخت T دارای می‌نیمم فاصله (و بنابراین در $C(T)$ است) است، اگر و تنها اگر $BW(v) \leq \frac{n-۱}{۲}$ باشد.

نتیجه بعد یکی از راه‌های محاسبه اندیس وینر را برای ما فراهم می‌آورد.

قضیه ۲-۳ فرض کنیم $e = (u, v)$ یک یال از گراف همبند G باشد. اگر $n_u(e)$ برابر تعداد رئوس از G که نزدیک‌تر به رأس u نسبت به رأس v باشد، $n_v(e)$ برابر تعداد رئوس از G که نزدیک‌تر به v نسبت به u باشد، آنگاه:

$$n_u(e) - n_v(e) = d(v) - d(u)$$

4) Jordan 5) Zelinka