

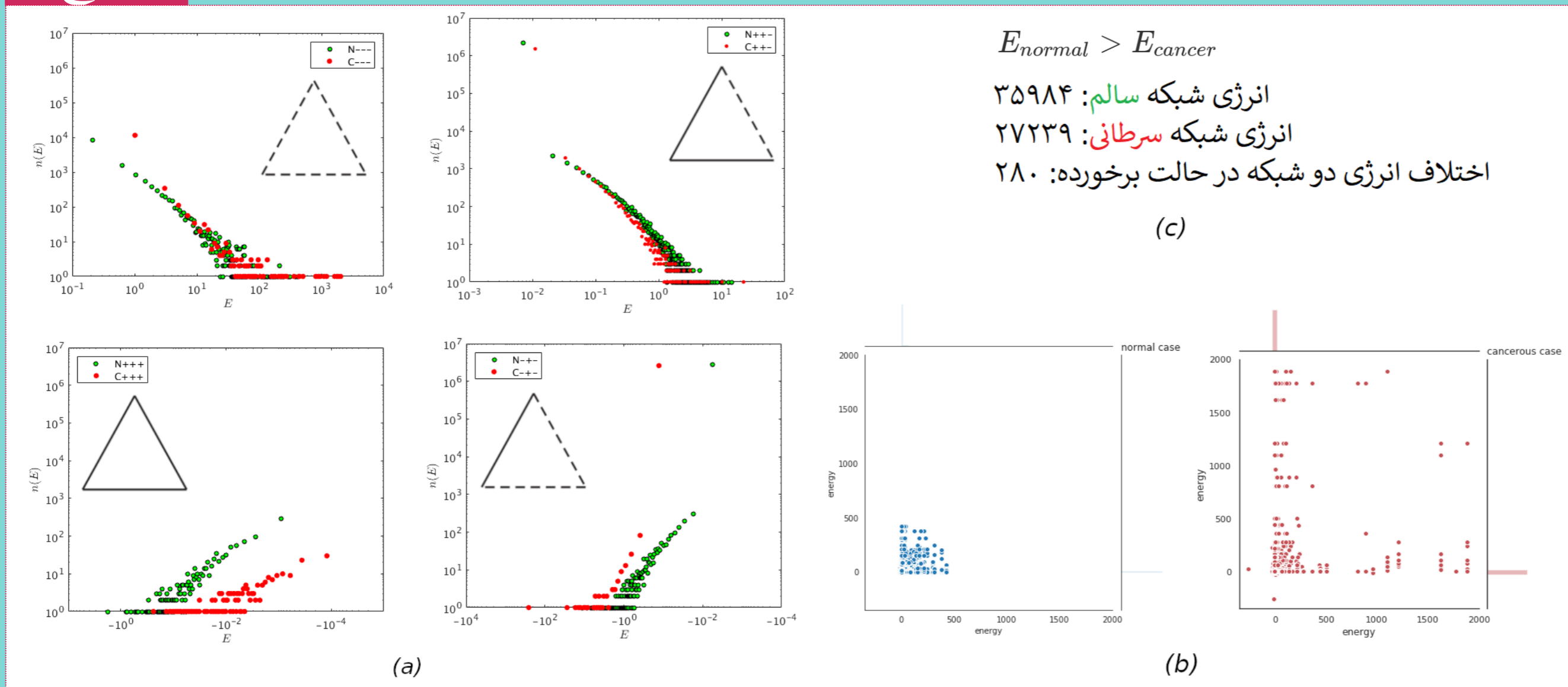


چشم‌انداز انرژی شبکه سلول سالم و سرطانی از افق نظریه توازن

عباس کریمی، مینا زمانی، امیرحسین شیرازی، غلامرضا جعفری

مرکز شبکه‌های پیچیده و علم داده اجتماعی
دانشگاه شهید بهشتی، دانشکده فیزیک

نتایج



خلاصه

نگاه غالب به سرطان در دهه‌های گذشته مبنی بر یافتن ژن‌های موثر و چشم‌پوشی از رفتار جمعی ژن‌ها درون سلول بوده است. در این مطالعه با استنباط شبکه تنظیمی ژن‌ها از روی داده‌های تجربی توالی‌یابی آرانی سرطان سینه و اعمال نظریه توازن، انرژی شبکه سالم و سرطانی به دست آمده است. نتایج حاکی از این است که انرژی شبکه سالم از شبکه سرطانی بیشتر است و توزیع مثلث‌های نامتوازن در شبکه سرطانی همگن‌تر از شبکه سالم است. از سوی دیگر مثلث‌های تشکیل شده در شبکه که واحدهای انرژی شبکه هستند، دارای توزیعی توانی می‌باشند که در شبکه سرطانی هم‌جورتر پخش شده‌اند.

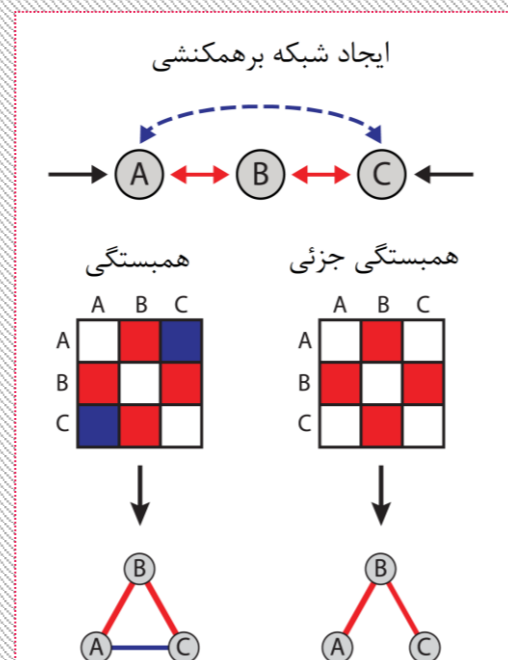
روش

هدف در فیزیک آماری وارون، استفاده از داده‌های موجود برای استنباط پارامترهایی است که یک مدل را توصیف می‌کنند، مثلاً یافتن ضرایب جفت‌شدگی در هامیلتونی یک سیستم اسپینی [۱]. برای این کار، روابط بین قوانین میکروسکوپی و مشاهده‌پذیرها باید معکوس شود. در مسئله شیشه-اسپین وارون، برهمکنش بین اسپین‌ها را نمی‌دانیم ولی می‌خواهیم با استفاده از اندازه‌گیری مغناطش‌ها، همبستگی‌ها و سایر مشاهده‌پذیرها آن‌ها را اصطلاحاً/استنباط کنیم. خوشبختانه روش‌های استنباط مبتنی بر آنتروپی بیشینه با موفقیت توانسته‌اند برهمکنش‌های مستقیم موجود در داده‌های زیستی را استنباط کنند [۲]. اکنون اگر فرض کنیم که هر ژن، یک اسپین و مقدار هر اسپین برابر با سطح بیان یک ژن است و مدل حاکم بر ژن‌ها یک مدل دوتایی مانند شیشه-اسپین است، روابط زیر برقرار خواهند بود؛ با حفظ عمومیت مسئله، دما را در دل جفت‌شدگی‌ها جا داده‌ایم و ثابت بولتزمن را برابر با یک گرفته‌ایم. توزیع زیر همان توزیعی است که آنتروپی گیبس-شانون S[P] را تحت شرایط بهنجار بودن $P(\{X_i\})$ و دارا بودن گشتاور اول و دوم، بیشینه می‌کند. از داده‌های تجربی مقادیر $X_i X_j$ قابل محاسبه است. این مقادیر می‌بایست برابر با گشتاور اول و دومی باشد که مدل ما تولید می‌کند. بنابراین با توجه به این قیود، بیشینه کردن آنتروپی توسط این مدل، ضرایب جفت‌شدگی یا همان پارامترهای مدل استنباط/یادگرفته می‌شوند. به سادگی نشان می‌توان داد [۲] که توزیع مورد نظر، یک توزیع چند متغیره گاوسی به شکل زیر است. با توجه به شکل تابع توزیع، ماتریس برهمکنش برابر خواهد بود با قرینه وارون ماتریس کواریانس. برای محاسبه وارون ماتریس کواریانس به خاطر تنگ بودن شبکه برهمکنشی از الگوریتم لاسو گرافیکی [۳] استفاده می‌کنیم. ماتریس C^{-1} اطلاعات مربوط به همبستگی جزئی متغیرها را نشان می‌دهد و J ماتریس همسایگی شبکه است. در علم شبکه، اطلاعات پیوندهای یک شبکه را به کمک ماتریس همسایگی آن نمایش می‌دهیم. از آن‌جا که درایه‌های ماتریس برهمکنش مقادیر حقیقی هستند، شبکه برهمکنشی ژن‌ها یک شبکه وزن‌دار علامت‌دار است. اکنون برای مطالعه سه‌تایی‌های این شبکه نظریه توازن هایدرا [۴] را اعمال می‌کنیم.

داده

در این مطالعه، مجموعه داده‌های بیان ۲۰۵۳۲ ژن از ۱۱۴ نمونه سالم و ۷۶۴ نمونه سرطانی از سرطان سینه به دست آمده از فناوری توالی‌یابی آرانی از پروژه TCGA به کار گرفته شده است. به منظور آن‌که مطالعه این دو دسته از داده‌ها از لحاظ محاسباتی مقدور باشد، از بین ۲۰۵۳۲ ژن، در نهایت ۴۱۳ ژن برای بررسی انتخاب شده است. از ژن‌ها و نمونه‌های موجود، ماتریسی با مقادیر بهنجار و استاندارد شده‌ای از مقادیر بیان شده ژن‌ها ساخته شده که از آن ماتریس کواریانس سطح بیان ژن‌ها قابل محاسبه است. هر درایه‌ای از ماتریس کواریانس، ضریب همبستگی پیرسون بین سطح بیان دو ژن متناظر است. می‌دانیم که همبستگی پیرسون بین دو متغیر شامل اطلاعات زایدی از تاثیر واسطه‌هاست، بنابراین برای حذف ارتباط واسطه‌ها و یافتن ارتباطات علی (برهمکنش‌ها) باید به سراغ همبستگی جزئی رفت.

$$H(X) = - \sum_{\langle ij \rangle} J_{ij} X_i X_j$$

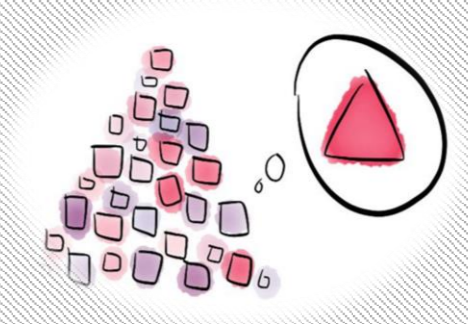


$$P(\mathbf{x}; \langle \mathbf{x} \rangle, C) = (2\pi)^{-L/2} \det(C)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \langle \mathbf{x} \rangle)^T C^{-1}(\mathbf{x} - \langle \mathbf{x} \rangle)\right)$$

$$J = -C^{-1}$$

$$E_{\Delta} = -J_{ij} J_{jk} J_{ki}$$

$$E_{Net} = \sum_{\Delta} E_{\Delta}$$



$$P(\{X_i\}) = \frac{1}{Z} \exp\left[\sum_{\langle ij \rangle} J_{ij} X_i X_j\right]$$

فیزیک

نتایج

از مشاهده توزیع مثلث‌های متوازن با انرژی منفی و مثلث‌های نامتوازن با انرژی مثبت، مشخص می‌شود که صرف نظر از رفتار توانی آن‌ها، شبکه برهمکنشی سالم دارای انرژی بیشتری نسبت به شبکه برهمکنشی سرطانی است که این به معنای پویایی بیشتر سلول سالم و توانایی آن در تغییر وضعیت خود است. اگر هر دو شبکه را بر بزنیم، اختلاف انرژی شبکه‌ها یک مرتبه بزرگی کم می‌شود. سلول سرطانی از آنجایی که انرژی کم‌تری دارد گویا در کمینه‌های موضعی پروفایل انرژی به دام می‌افتد و به همین خاطر به دلیل پویا نبودن دینامیکی آن گویا کاری جز باز تولید ندارد. حرکت جمعی یک سیستم، متناظر با دوتایی‌های موجود در آن سیستم است.

در این مقاله نشان داده شد که برای مطالعه سرطان، فقط رفتار تک ژن‌ها در سرشت سلول مهم نیست و ساختارهای مرتبه-بالا نیز در تشکیل رفتار جمعی دخیل هستند.

- شبکه سالم نسبت به شبکه سرطانی تمایل بیشتری به تغییر وضعیت پیوندهای خود دارد. (فعالیت‌های تنظیمی مختلف)
- تعداد مثلث‌های نامتوازن در دو حالت تقریباً برابر است ولی توزیع آن‌ها در شبکه سالم و سرطانی بسیار متفاوت است.
- تعداد مثلث‌های متوازن با سه پیوند مثبت در شبکه سالم بیشتر از شبکه سرطانی است و تعداد مثلث‌های متوازن با یک پیوند مثبت برعکس.

نظریه توازن

اگر دوستی بین دو شخص i و j را با $J_{ij} > 0$ و دشمنی بین آن دو را با $J_{ij} < 0$ نشان دهیم، آن‌گاه برای هر مثلث، چهار حالت موجود است: (۱) دوست دوست شما، دوست شماست. (۲) دشمن دوست شما، دشمن شماست. (۳) دوست دشمن شما، دشمن شماست. (۴) دشمن دشمن شما، دوست شماست. مقدار J_{ij} اندازه رابطه دوستی و دشمنی را مشخص می‌کند که عددی حقیقی است. بنابراین شبکه تنظیم ژن، یک شبکه وزن‌دار است. اکنون می‌توانیم به هر مثلث انرژی E_{Δ} را نسبت دهیم و با جمع زدن روی همه مثلث‌های شبکه، انرژی آن شبکه را محاسبه کنیم [۵]. مثلث‌های متوازن دارای انرژی منفی، مشابه با مفهوم پایداری در وضعیت دینامیکی یا مفهوم ناکامی (فراسریش) در فیزیک و مثلث‌های نامتوازن دارای انرژی مثبت، در تقابل با پایداری اجتماعی و تمایل به تغییر وضعیت هستند. توزیع مثلث‌های نامتوازن و متوازن در بالای صفحه آمده است.

References:
[1] Michele Castellana, William Bialek - Inverse spin glass and related maximum entropy problems - Physical Review Letters 113(11), 117204 (2014)
[2]: Stein RR, Marks DS, Sander C (2015) Inferring Pairwise Interactions from Biological Data Using Maximum-Entropy Probability Models. PLoS Comput Biol 11(7): e1004182.

[3] Friedman, Jerome and Hastie, Trevor and Tibshirani, Robert (2008). "Sparse inverse covariance estimation with the graphical lasso" Biostatistics. Biometrika Trust.
[5] Heider, Fritz (1958). The Psychology of Interpersonal Relations. John Wiley & Sons.