



Aprendizado Semi-Supervisionado em Redes Complexas

Fabricio Breve
ICMC – USP

Conteúdo

- Introdução
- Aprendizado Semi-Supervisionado
- Classes de Algoritmos
 - Modelos Baseados em Grafos
 - Construção do Grafo
 - Métodos Iterativos
 - Regularização
 - Modelo de Partículas

Introdução

- Redes a serem analisadas estão se tornando cada vez maiores
 - Internet
 - World Wide Web (WWW)
 - Sistemas de telecomunicações
 - Redes de energia elétrica
 - Redes sociais
 - Redes de tráfego
 - Redes biológicas
 - Redes neurais
 - Redes de interação entre proteínas

Introdução

- Freqüentemente o processo de rotular dados é:
 - Caro
 - Demorado
 - Requer o trabalho de um especialista humano
- Conseqüência: Em muitas situações, apenas um pequeno subconjunto de itens pode ser rotulado

Introdução

- **Aprendizado Supervisionado**
 - Apenas amostras rotuladas são usadas para o treinamento
- **Aprendizado Não Supervisionado**
 - Todas as amostras são não rotuladas
- **Aprendizado Semi-Supervisionado**
 - Combina uma pequena quantidade de amostras rotuladas com um grande número de amostras não rotuladas para produzir melhores classificadores

Aprendizado Semi-Supervisionado

- Tipicamente o conjunto de dados X pode ser dividido em duas partes:
 - $X_l = (x_1, \dots, x_l)$, para qual os rótulos $Y_l = (y_1, \dots, y_l)$ são fornecidos
 - $X_u = (x_{l+1}, \dots, x_{l+u})$, para qual os rótulos são desconhecidos

Aprendizado Semi-Supervisionado

- Quando funciona?
 - Comparando com um algoritmo supervisionado que usa apenas dados rotulados, podemos esperar melhores resultados considerando as amostras não rotuladas?
 - Sim, desde que a distribuição dos dados (a ser revelada) seja relevante para o problema de classificação.

Aprendizado Semi-Supervisionado

- Para funcionar algumas hipóteses precisam ser verdadeiras
 - Hipótese da suavidade
 - Se dois pontos x_1, x_2 em uma região de alta-densidade estão próximos, então suas saídas y_1, y_2 também deverão ser próximas.
 - Hipótese de cluster
 - Se dois pontos estão no mesmo cluster, então provavelmente eles são da mesma classe.
 - Não implica que cada classe é apenas um cluster
 - Pode ser considerado um caso especial da hipótese da suavidade

Aprendizado Semi-Supervisionado

- Algumas aplicações práticas:
 - Em reconhecimento de fala, custa quase nada gravar grandes quantidades de dados, mas rotulá-los requer que um humano ouça e digite uma transcrição.
 - Bilhões de páginas estão diretamente disponíveis para processamento direto, mas para classificá-las humanos tem de lê-las.
 - Seqüências de proteínas são fáceis de adquirir em velocidade industrial hoje em dia (por seqüenciamento de genoma, busca computacional de gene, e tradução automática), mas resolver uma estrutura tridimensional ou determinar as funções de uma simples proteína pode levar anos de trabalho científico.

Classes de Algoritmos

- Auto Treinamento
- Modelos generativos
- Separação de baixa densidade
- **Métodos baseados em grafos**

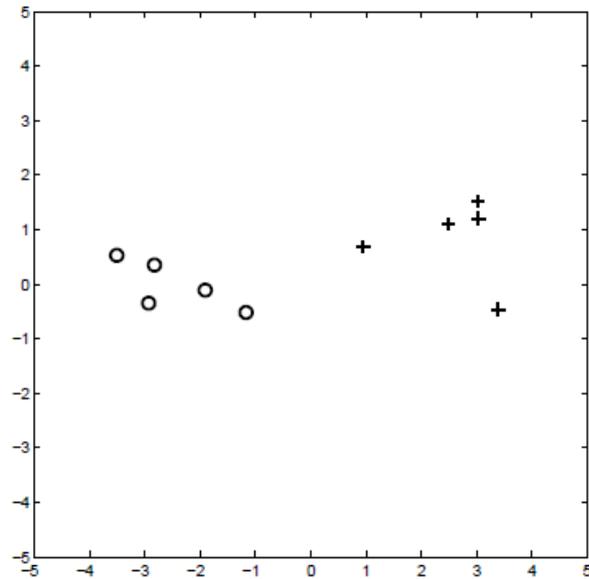
- X. Zhu, "Semi-supervised learning literature survey," Computer Sciences, University of Wisconsin-Madison, Tech. Rep. 1530, 2005.
- O. Chapelle, B. Schölkopf, and A. Zien, Eds., Semi-Supervised Learning, ser. Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press, 2006.

Auto Treinamento

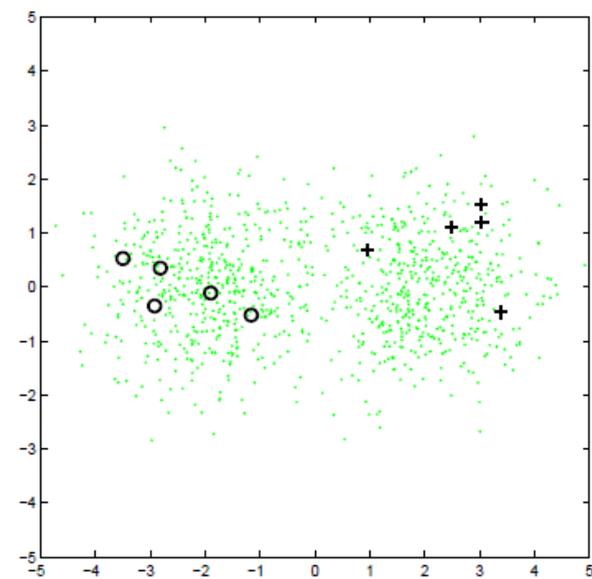
1. Um classificador é treinado usando os poucos dados rotulados.
 2. Esse classificador é usado para classificar os dados não rotulados.
 3. Os dados não rotulados mais confiáveis e seus rótulos preditos são adicionados ao conjunto de treinamento.
 4. O classificador é re-treinado e o procedimento é repetido.
- Também chamado *auto aprendizado* ou *bootstrapping*
 - Alguns algoritmos evitam que erros de classificação sejam reforçados “desaprendendo” dados rotulados cuja confiança caia abaixo de um certo nível.

Modelos Generativos

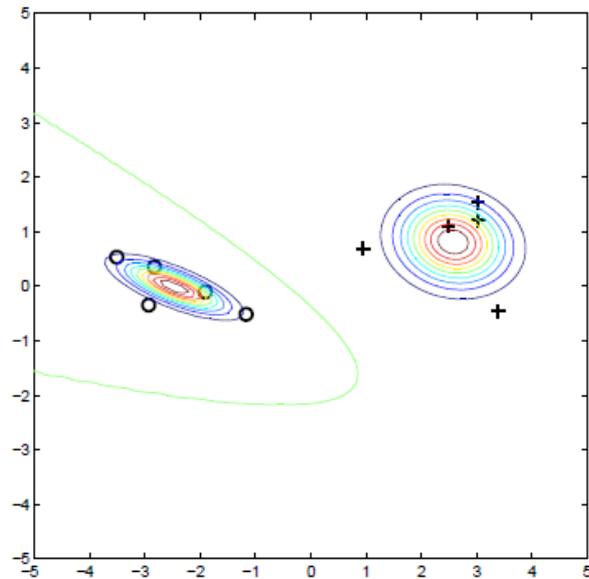
- Assumem um modelo $p(x,y) = p(y) p(x|y)$ onde $p(x|y)$ é uma distribuição identificável, por exemplo Gaussiana
- Com um grande número de dados não rotulados a distribuição pode ser identificada; então idealmente precisaríamos apenas de uma amostra rotulada por componente para determinar a distribuição



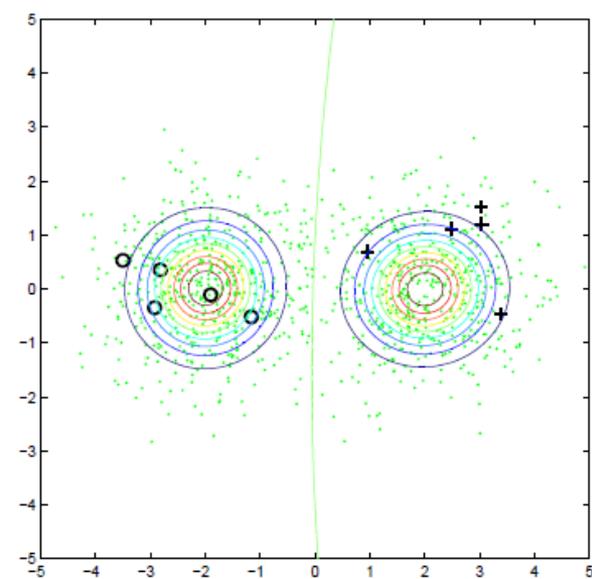
(a) dados rotulados



(b) dados rotulados e não rotulados (pontos pequenos)



(c) modelo aprendido dos dados rotulados



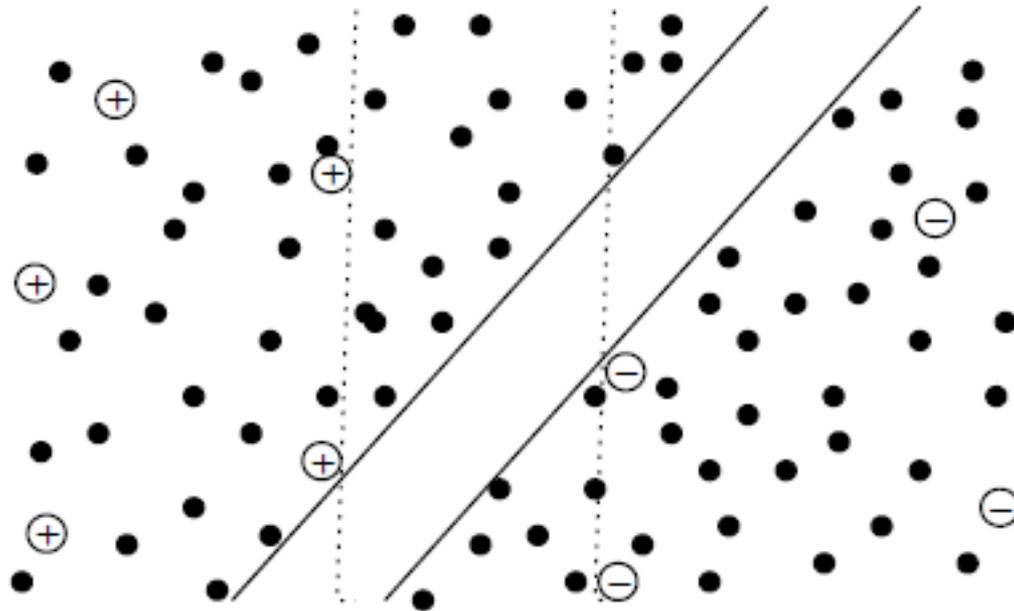
(d) modelo aprendido dos dados rotulados e não rotulados

X. Zhu, "Semi-supervised learning literature survey," Computer Sciences, University of Wisconsin-Madison, Tech. Rep. 1530, 2005.

Problema de classificação binária, se assumimos que cada classe tem distribuição Gaussiana, então podemos usar dados não-rotulados para ajudar a estimar os parâmetros

Separação de Baixa Densidade

- Tentam empurrar a fronteira de decisão para longe dos dados não rotulados
- Exemplo:
 - Transductive Support Vector Machines (TSVM)



Métodos Baseados em Grafos

- Área de pesquisa mais ativa em aprendizado semi-supervisionado.
- Dados são representados por nós de um grafo, onde as arestas são rotuladas com a distância entre os pares de nós.
- Nós rotulados são usados para propagar informação de rótulo para os demais.
- Métodos assumem suavidade de rótulos através do grafo.

Métodos Baseados em Grafos

- X. Zhu, Z. Ghahramani, and J. Lafferty, “Semi-supervised learning using gaussian fields and harmonic functions,” in *Proceedings of the Twentieth International Conference on Machine Learning*, 2003, pp. 912–919.
- D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, “Learning with local and global consistency,” in *Advances in Neural Information Processing Systems*, vol. 16. MIT Press, 2004, pp. 321–328. [Online]. Available: <http://www.kyb.tuebingen.mpg.de/bs/people/weston/localglobal.pdf>
- X. Zhu and Z. Ghahramani, “Learning from labeled and unlabeled data with label propagation,” Carnegie Mellon University, Pittsburgh, Tech. Rep. CMU-CALD-02-107, 2002. [Online]. Available: <http://citeseer.ist.psu.edu/581346.html>
- F. Wang and C. Zhang, “Label propagation through linear neighborhoods,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 1, pp. 55–67, Jan. 2008.
- A. Blum and S. Chawla, “Learning from labeled and unlabeled data using graph mincuts,” in *Proceedings of the Eighteenth International Conference on Machine Learning*. San Francisco: Morgan Kaufmann, 2001, pp. 19–26.
- M. Belkin, I. Matveeva, and P. Niyogi, “Regularization and semisupervised learning on large graphs,” in *Conference on Learning Theory*. Springer, 2004, pp. 624–638.
- M. Belkin, N. P., and V. Sindhwani, “On manifold regularization,” in *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTAT 2005)*. New Jersey: Society for Artificial Intelligence and Statistics, 2005, pp. 17–24.
- T. Joachims, “Transductive learning via spectral graph partitioning,” in *Proceedings of International Conference on Machine Learning*. AAAI Press, 2003, pp. 290–297.

Métodos Baseados em Grafos

- Seja um grafo $G = (V, E)$, com $V = \{v_1, v_2, \dots, v_n\}$, onde cada nó v_i corresponde a uma amostra x_i
- Uma matriz de adjacência W define quais nós da rede estão interconectados, ou seja, ele identifica os nós em E

$$W_{ij} = \begin{cases} w_{ij}(e) & \text{se } i \text{ e } j \text{ são vizinhos} \\ 0 & \text{se } i \text{ e } j \text{ não são vizinhos} \end{cases}$$

$w_{ij}(e)$ pode ser um número real que mede a similaridade entre i e j (por exemplo)

Métodos de Construção do Grafo

- A matriz de pesos W pode ser dada pelos k -vizinhos mais próximos
 - $W_{ij} = 1$ se x_i está entre os k -vizinhos mais próximos de x_j ou vice-versa (e 0 caso contrário)
- Outra matriz de peso típica é dada pelo kernel Gaussiano de largura σ

$$\mathbf{w}_{ij} = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (11.1)$$

Métodos Iterativos

- Nós rotulados iniciam com seus respectivos rótulos (1 ou -1)
- Nós não rotulados iniciam com 0
- Nós propagam seus rótulos repetidamente até convergência
 - Exemplo: X. Zhu and Z. Ghahramani, “Learning from labeled and unlabeled data with label propagation,” Carnegie Mellon University, Pittsburgh, Tech. Rep. CMU-CALD-02-107, 2002. [Online]. Available: <http://citeseer.ist.psu.edu/581346.html>

Métodos Iterativos

Algorithm 11.1 Label propagation (Zhu and Ghahramani, 2002)

Compute affinity matrix \mathbf{W} from (11.1)

Compute the diagonal degree matrix \mathbf{D} by $\mathbf{D}_{ii} \leftarrow \sum_j W_{ij}$

Initialize $\hat{Y}^{(0)} \leftarrow (y_1, \dots, y_l, 0, 0, \dots, 0)$

Iterate

1. $\hat{Y}^{(t+1)} \leftarrow \mathbf{D}^{-1} \mathbf{W} \hat{Y}^{(t)}$

2. $\hat{Y}_l^{(t+1)} \leftarrow Y_l$

until convergence to $\hat{Y}^{(\infty)}$

Label point x_i by the sign of $\hat{y}_i^{(\infty)}$

- Alternativas:
 - Forçar $\mathbf{W}_{ii} = 0$ pode resultar em melhores resultados
 - Remover a linha 2 pode dar melhores resultados quando as classes se sobrepõem

Métodos Iterativos

- D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, “Learning with local and global consistency,” in *Advances in Neural Information Processing Systems*, vol. 16. MIT Press, 2004, pp. 321–328.
 - A cada passo um nó i recebe a contribuição de seus vizinhos j (ponderados pelo peso normalizado da aresta (i,j)), e uma pequena contribuição adicional de seu valor inicial

Algorithm 11.3 Label spreading (Zhou et al., 2004)

Compute the affinity matrix \mathbf{W} from (11.1) for $i \neq j$ (and $\mathbf{W}_{ii} \leftarrow 0$)

Compute the diagonal degree matrix \mathbf{D} by $\mathbf{D}_{ii} \leftarrow \sum_j W_{ij}$

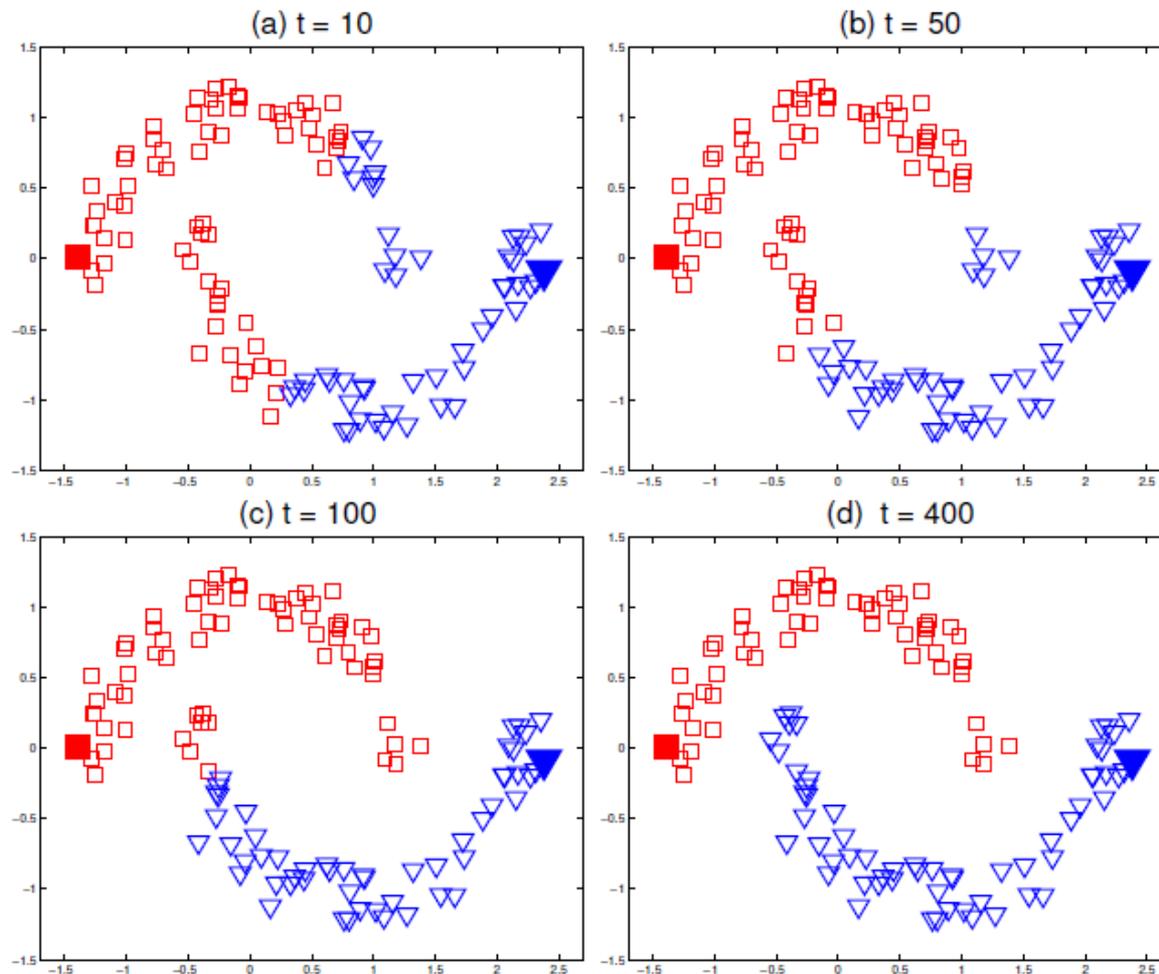
Compute the normalized graph Laplacian $\mathcal{L} \leftarrow \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$

Initialize $\hat{Y}^{(0)} \leftarrow (y_1, \dots, y_l, 0, 0, \dots, 0)$

Choose a parameter $\alpha \in [0, 1)$

Iterate $\hat{Y}^{(t+1)} \leftarrow \alpha \mathcal{L} \hat{Y}^{(t)} + (1 - \alpha) \hat{Y}^{(0)}$ until convergence to $\hat{Y}^{(\infty)}$

Label point x_i by the sign of $\hat{y}_i^{(\infty)}$



Classificação de padrão com duas luas dado pelo modelo de Zhou et. al. (2004). O processo de convergência do algoritmo com t crescendo de 1 a 400 é mostrado de (a) até (d) Note que as informações de rótulos inicial é difundida ao longo das luas.

Regularização em Grafos

- O problema do aprendizado semi-supervisionado em um grafo \mathbf{G} consiste em encontrar um conjunto de rótulos que seja consistente com ambos o conjunto de rótulos iniciais (incompleto) e a geometria dos dados induzida pela estrutura do grafo (arestas e pesos em \mathbf{W})

Regularização em Grafos

- Consistência com os rótulos iniciais:

$$\sum_{i=1}^l (\hat{y}_i - y_i)^2 = \|\hat{Y}_l - Y_l\|^2$$

- Consistência com a geometria dos dados

$$\begin{aligned} \frac{1}{2} \sum_{i,j=1}^n \mathbf{W}_{ij} (\hat{y}_i - \hat{y}_j)^2 &= \frac{1}{2} \left(2 \sum_{i=1}^n \hat{y}_i^2 \sum_{j=1}^n \mathbf{W}_{ij} - 2 \sum_{i,j=1}^n \mathbf{W}_{ij} \hat{y}_i \hat{y}_j \right) \\ &= \hat{Y}^\top (\mathbf{D} - \mathbf{W}) \hat{Y} \\ &= \hat{Y}^\top L \hat{Y} \end{aligned}$$

com $L = \mathbf{D} - \mathbf{W}$. Isto significa que penalizamos mudanças rápidas em \hat{Y} entre pontos que estão próximos (dado pela matriz de similaridade \mathbf{W})

Regularização em Grafos

- Vários algoritmos são baseados nessas considerações
 - X. Zhu, Z. Ghahramani, and J. Lafferty. Semi-supervised learning using Gaussian fields and harmonic functions. In Twentieth International Conference on Machine Learning, pages 912–912, Washington, DC, 2003b. AAAI Press.
 - Força os rótulos iniciais dos dados rotulados ($\hat{Y}_l = Y_l$).
 - Então minimiza a função de consistência da geometria em \hat{Y}_u

Regularização em Grafos

- Porém, quando há ruído nos rótulos disponíveis, pode ser benéfico permitir que os dados rotulados sejam re-rotulados.
- Também pode melhorar a generalização mesmo quando não há ruído
- Isto leva a um critério de custo mais geral, envolvendo um balanço entre as duas equações.
 - M. Belkin, I. Matveeva, and P. Niyogi. Regularization and semi-supervised learning on large graphs. In Proceedings of the Seventeenth Annual Conference on Computational Learning Theory, pages 624–638, Banff, Canada, 2004b.
 - O. Delalleau, Y. Bengio, and N. Le Roux. Efficient non-parametric function induction in semisupervised learning. In Artificial Intelligence and Statistics, 2005.
- Um termo de regularização pode ser adicionado

$$C(\hat{Y}) = \|\hat{Y}_l - Y_l\|^2 + \mu \hat{Y}^\top L \hat{Y} + \mu \epsilon \|\hat{Y}\|^2$$

Regularização em Grafos

- A maioria dos métodos iterativos também pode ser visto como estimando uma função f no grafo que satisfaça as seguintes condições ao mesmo tempo:
 1. Deve ser próximo dos rótulos dados pelo nós pré-rotulados
 2. Deve ser suave por todo o grafo.
- Isto pode ser visto como um framework de regularização onde o primeiro termo é uma função de perda, e o segundo termo é um regularizador.

Regularização em Grafos

- Exemplo:
 - O algoritmo “Label Spreading” (ou Método de Consistência) equivale à seguinte função de custo:

$$Q(F) = \frac{1}{2} \left(\sum_{i,j=1}^n W_{ij} \left\| \frac{1}{\sqrt{D_{ii}}} F_i - \frac{1}{\sqrt{D_{jj}}} F_j \right\|^2 + \mu \sum_{i=1}^n \|F_i - Y_i\|^2 \right)$$

- Onde $\mu > 0$ é o parâmetro de regularização. Então a função de classificação é:

$$F^* = \arg \min_{F \in \mathcal{F}} Q(F)$$

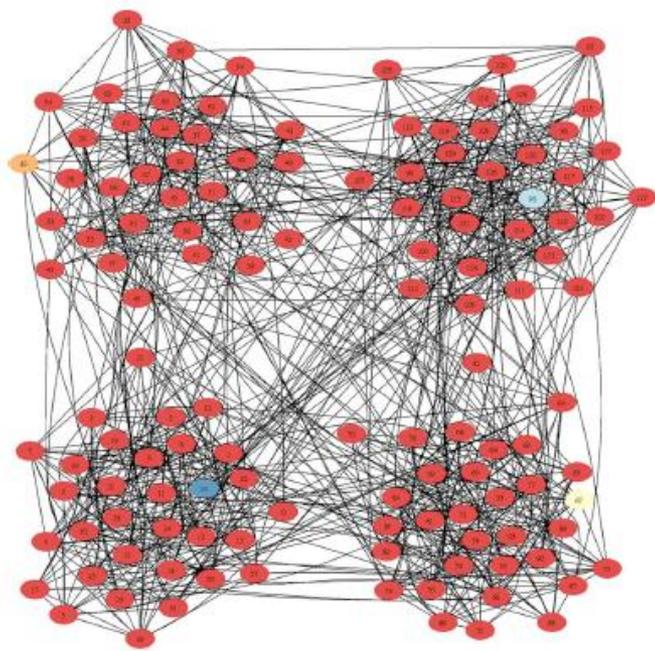
- D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, “Learning with local and global consistency,” in *Advances in Neural Information Processing Systems*, vol. 16. MIT Press, 2004, pp. 321–328.

Métodos Baseados em Grafos

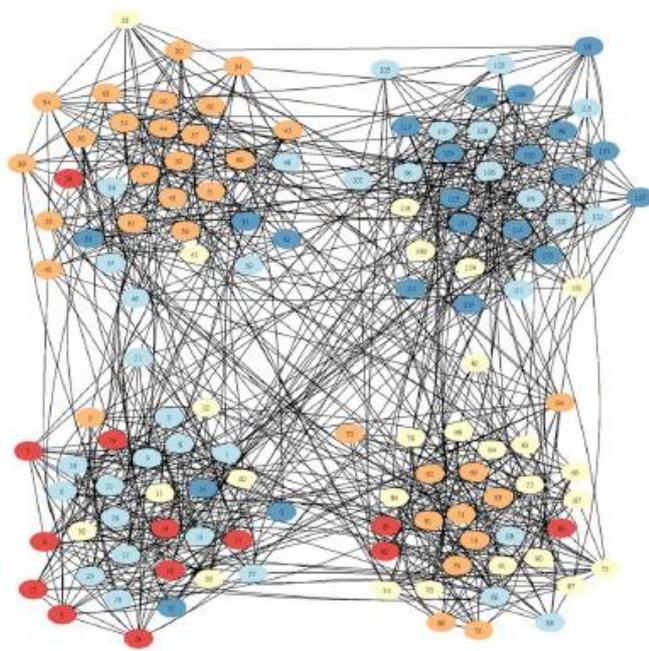
- Podem identificar diferentes distribuições de dados
- A maioria deles compartilha de um framework de regularização, diferindo apenas na escolha particular da função de perda e do regularizador
- A maioria deles tem alta ordem de complexidade computacional ($O(n^3)$), limitando seu uso a base de dados pequenas ou médias.

Modelo de Partículas

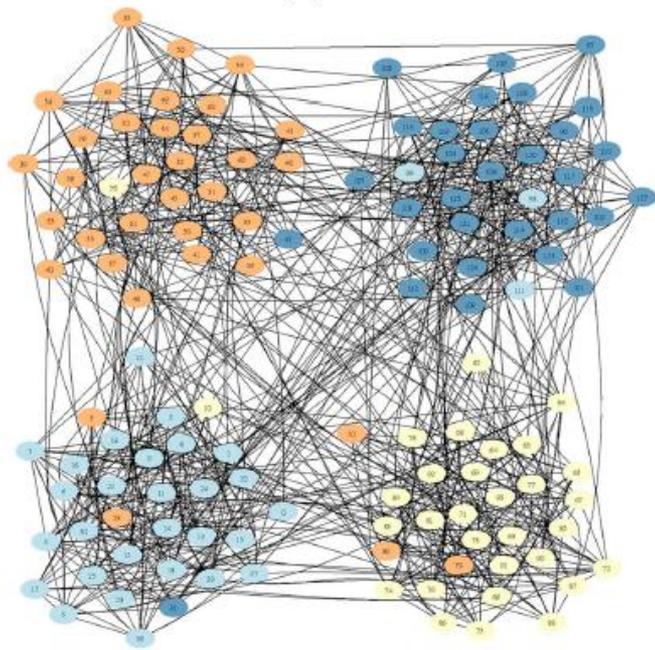
- M. G. Quiles, L. Zhao, R. L. Alonso, and R. A. F. Romero, “Particle competition for complex network community detection,” *Chaos*, vol. 18, no. 3, p. 033107, 2008. [Online]. Available: <http://link.aip.org/link/?CHAOEH/18/033107/1>
 - Partículas caminham na rede e competem com as outras de forma que cada uma tenta possuir a maior quantidade de nós possível
 - Cada partícula tenta evitar que outras partículas invadam seu território
 - Finalmente, cada partícula é confinada dentro de uma comunidade na rede



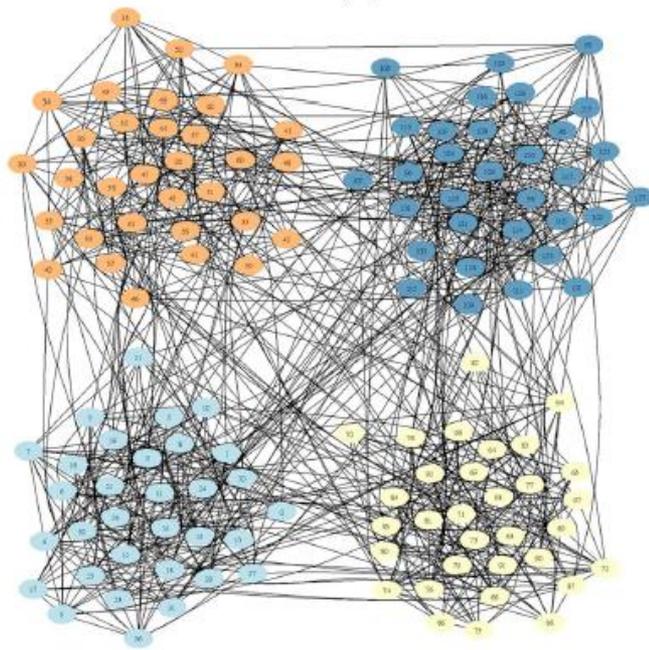
(a)



(b)



(c)



(d)

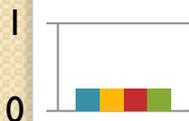
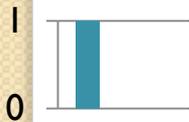
Ilustração do processo de detecção de comunidade pela caminhada competitiva de partículas competitiva. O número total de nós é $N=128$, o número de comunidades é $M=4$. A proporção de links externos é $z_{out} / k = 0.2$, e o grau médio dos nós é $k=16$. (a)

Configuração inicial.

Quatro partículas representadas por amarelo, ciano, laranja e azul, são colocadas aleatoriamente na rede. Vermelho representa nós livres. (b) Um snapshot da iteração 250. (c) Um snapshot da iteração 3500. (d) Um snapshot da iteração 7000.

Configuração Inicial

- Uma partícula é gerada para cada nó rotulado na rede
 - Nó-casa da partícula
- Partículas com o mesmo rótulo jogam pelo mesmo time
- Nós tem um vetor de domínio
 - Nós rotulados tem o domínio configurado para seus respectivos times
 - Ex: [1 0 0 0] (4 classes, nós rotulados como classe A)
 - Nós não rotulados tem níveis configurados igualmente para cada time
 - Ex: [0.25 0.25 0.25 0.25] (4 classes, nós não rotulados)
- Posição inicial das partículas é configurada para seus respectivos nós-casa

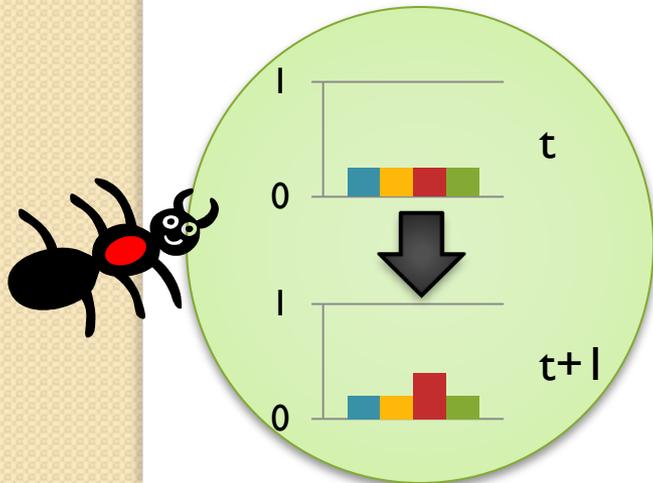


Modelo de Partículas Semi-Supervisionado

- Competição e cooperação entre partículas na rede
 - Competição pela posse de nós da rede
 - Cooperação entre partículas do mesmo time (rótulo)
 - Cada time de partículas tenta dominar a maior quantidade de nós possível de maneira cooperativa e ao mesmo tempo evitar a invasão de partículas de outros times.
 - Caminhada Aleatória-Determinística

Dinâmica de nós

- Quando uma partícula seleciona um vizinho para visitar:
 - Ela diminui o nível de domínio de outros times no nó alvo.
 - Ela aumenta o nível de domínio de seu próprio time no nó alvo.
 - Exceção:
 - Níveis de domínio de nós rotulados é fixo.



$$v_i^{\omega_\ell}(t+1) = \begin{cases} \max\left\{0, v_i^{\omega_\ell}(t) - \frac{\Delta v \rho_j^\omega(t)}{c-1}\right\} & \text{if } y_i \notin L \text{ and } \ell \neq \rho_j^f \\ v_i^{\omega_\ell}(t) + \sum_{q \neq \ell} v_i^{\omega_q}(t) - v_i^{\omega_q}(t+1) & \text{if } y_i \notin L \text{ and } \ell = \rho_j^f \\ v_i^{\omega_\ell}(t) & \text{if } y_i \in L \end{cases}$$

Dinâmica de Partículas

- Uma partícula se torna:
 - Mais forte quando ela tem como alvo um nó dominado por seu time
 - Mais fraca quando ela tem como alvo um nó dominado por outro time

$$\rho_j^\omega(t+1) = \rho_j^\omega(t) + \Delta_\rho(v_i^{\omega\ell}(t+1) - \rho_j^\omega(t))$$

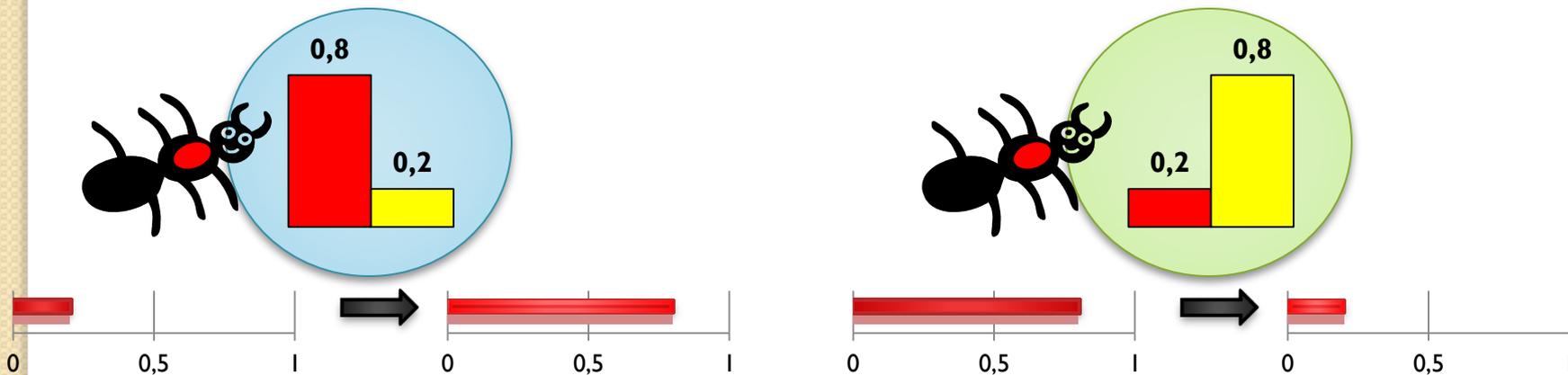
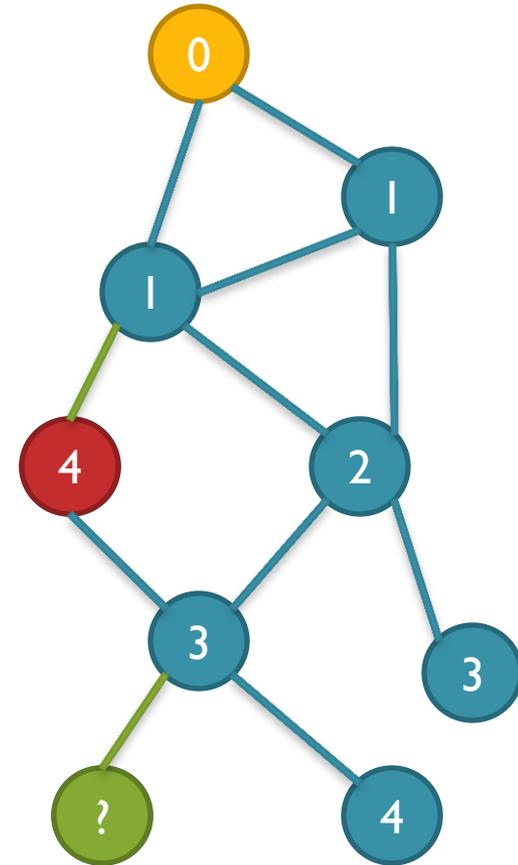


Tabela de Distância

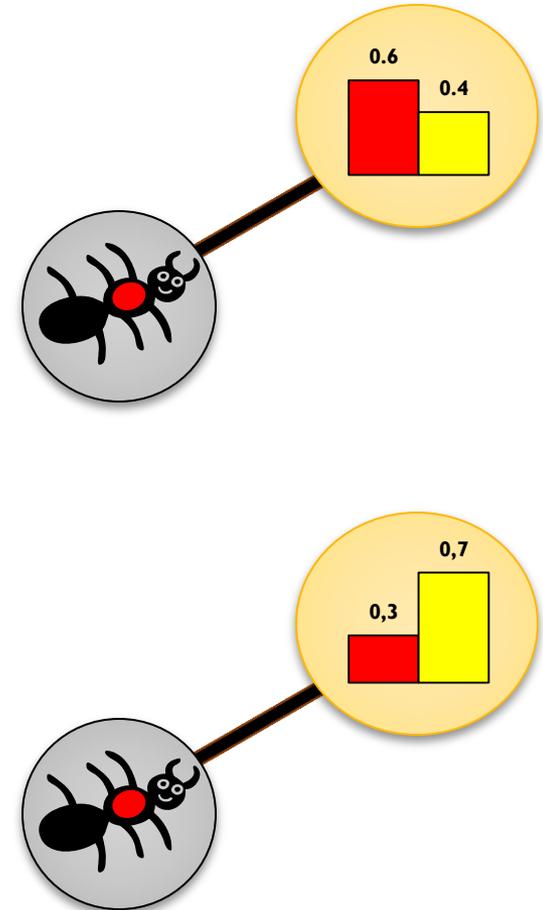
- Mantém a partícula informada da distância de seu nó-casa
 - Evita que a partícula perca toda sua força quando andando em vizinhanças inimigas
 - Mantém as partículas por perto para proteger sua própria vizinhança
- Atualizadas dinamicamente com informação local
 - Não requer cálculo a priori



$$\rho_j^{d_k}(t+1) = \begin{cases} \rho_j^{d_i}(t) + 1 & \text{if } \rho_j^{d_i}(t) + 1 < \rho_j^{d_k}(t) \\ \rho_j^{d_k}(t) & \text{otherwise} \end{cases}$$

Caminhada de partículas

- Choques
 - Uma partícula visita o nó alvo somente se o nível de domínio de seu time é maior que o dos outros;
 - Caso contrário, um choque acontece e a partícula fica no nó onde estava até a próxima iteração.
- Como uma partícula escolhe um vizinho como alvo?
 - Caminhada aleatória
 - Caminhada determinística

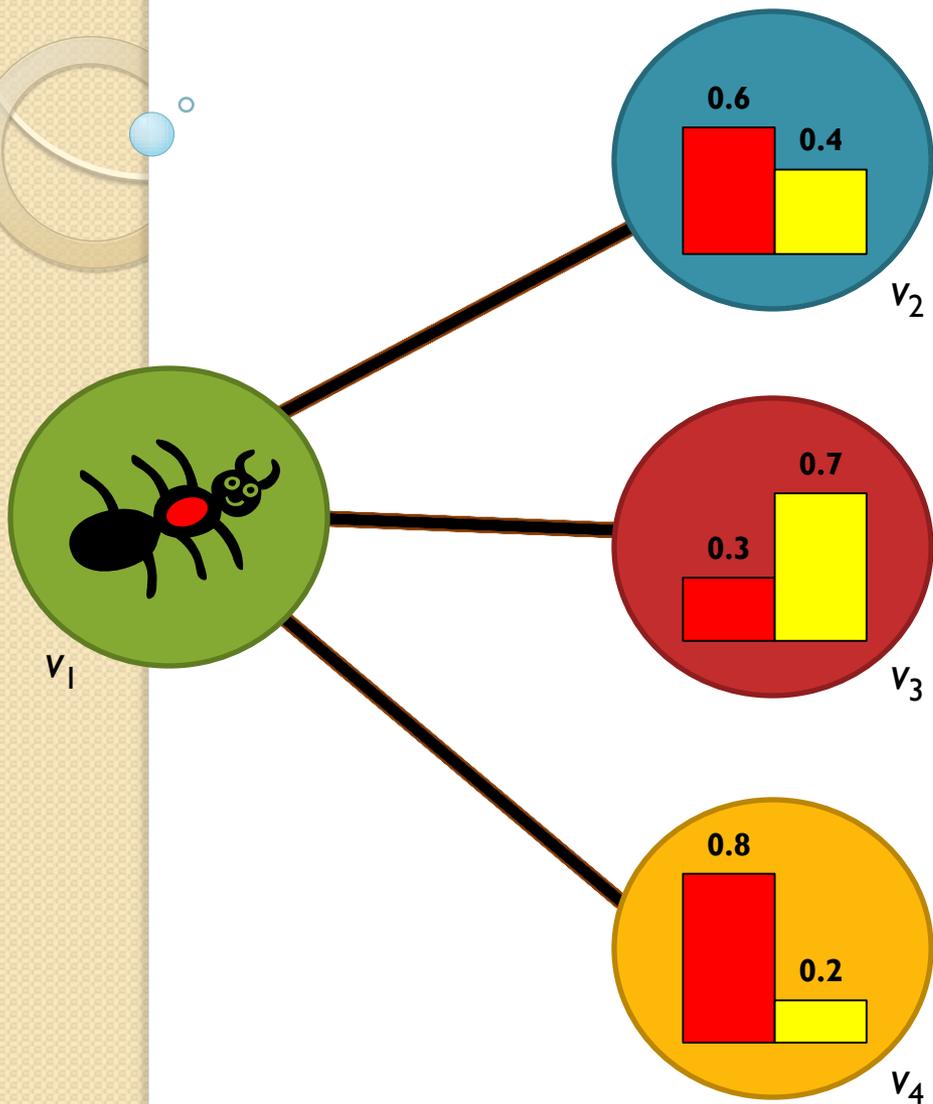


Caminhada Aleatória-Determinística

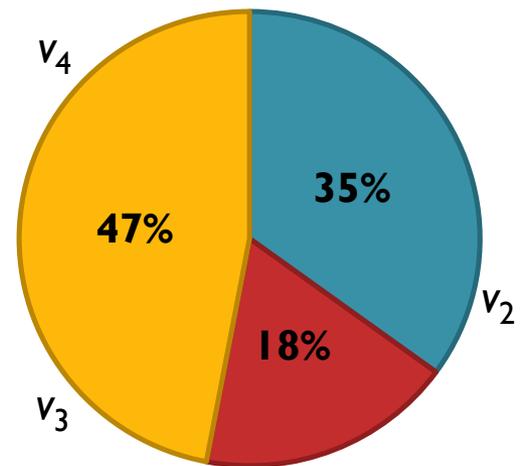
- Caminhada Aleatória
 - A partícula escolhe aleatoriamente um vizinho para visitar sem se preocupar com níveis de domínio ou distância
- Caminhada Determinística
 - A partícula prefere visitar nós que seu time já domina e nós próximos de sua casa

$$p(v_i|\rho_j) = \frac{W_{ki}}{\sum_{q=1}^n W_{qi}} \quad p(v_i|\rho_j) = \frac{W_{ki}v_i^{\omega_\ell} \frac{1}{(1+\rho_j^{d_i})^2}}{\sum_{q=1}^n W_{qi}v_i^{\omega_\ell} \frac{1}{(1+\rho_j^{d_i})^2}}$$

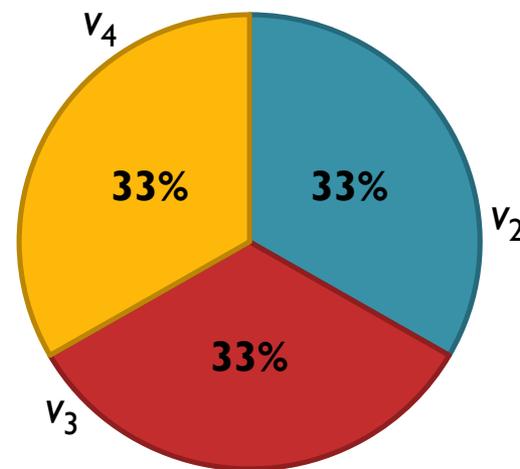
A partícula precisa exibir ambos os movimentos para obter um equilíbrio entre comportamento exploratório e defensivo



Probabilidades no Movimento Determinístico

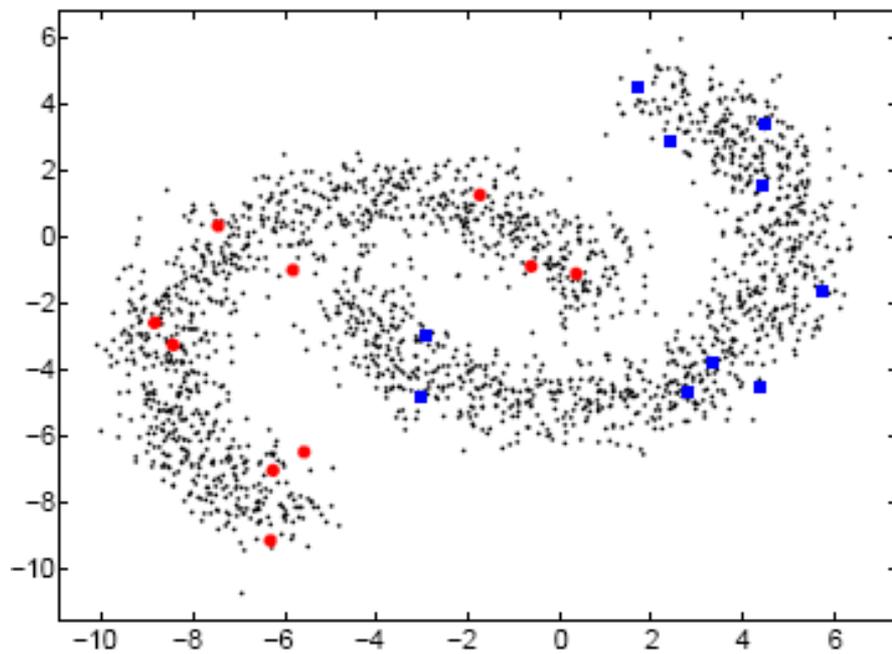


Probabilidades no Movimento Aleatório

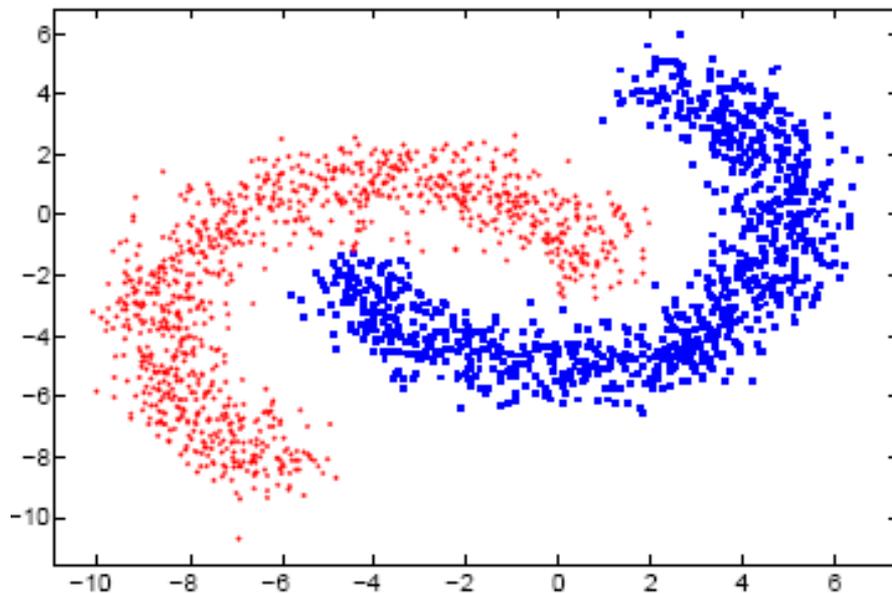


Algoritmo

- 1) Construir a matriz de adjacência,
- 2) Configurar os níveis de domínio,
- 3) Configurar posição inicial de partículas e seus correspondentes nós-casa. Configurar força da partícula e distância,
- 4) Repetir passos 5 a 9 até a convergência ou até que um número pré-definido de passos seja atingido,
- 5) Para cada partícula, complete os passos 6 a 9,
- 6) Selecione o nó alvo usando a regra determinística-aleatória,
- 7) Atualize os níveis de domínio do nó alvo,
- 8) Atualize a força da partícula,
- 9) Atualize a tabela de distâncias,
- 10) Rotule cada nó não rotulado de acordo com o time que tiver maior domínio.



(a)



(b)

Fig. 1. Classificação de padrões em forma de banana. (a) base de dados com 2.000 elementos divididos em duas classes, 20 amostras são pré-rotuladas (círculos vermelhos e quadrados azuis). (b) classificação obtida pelo método proposto.

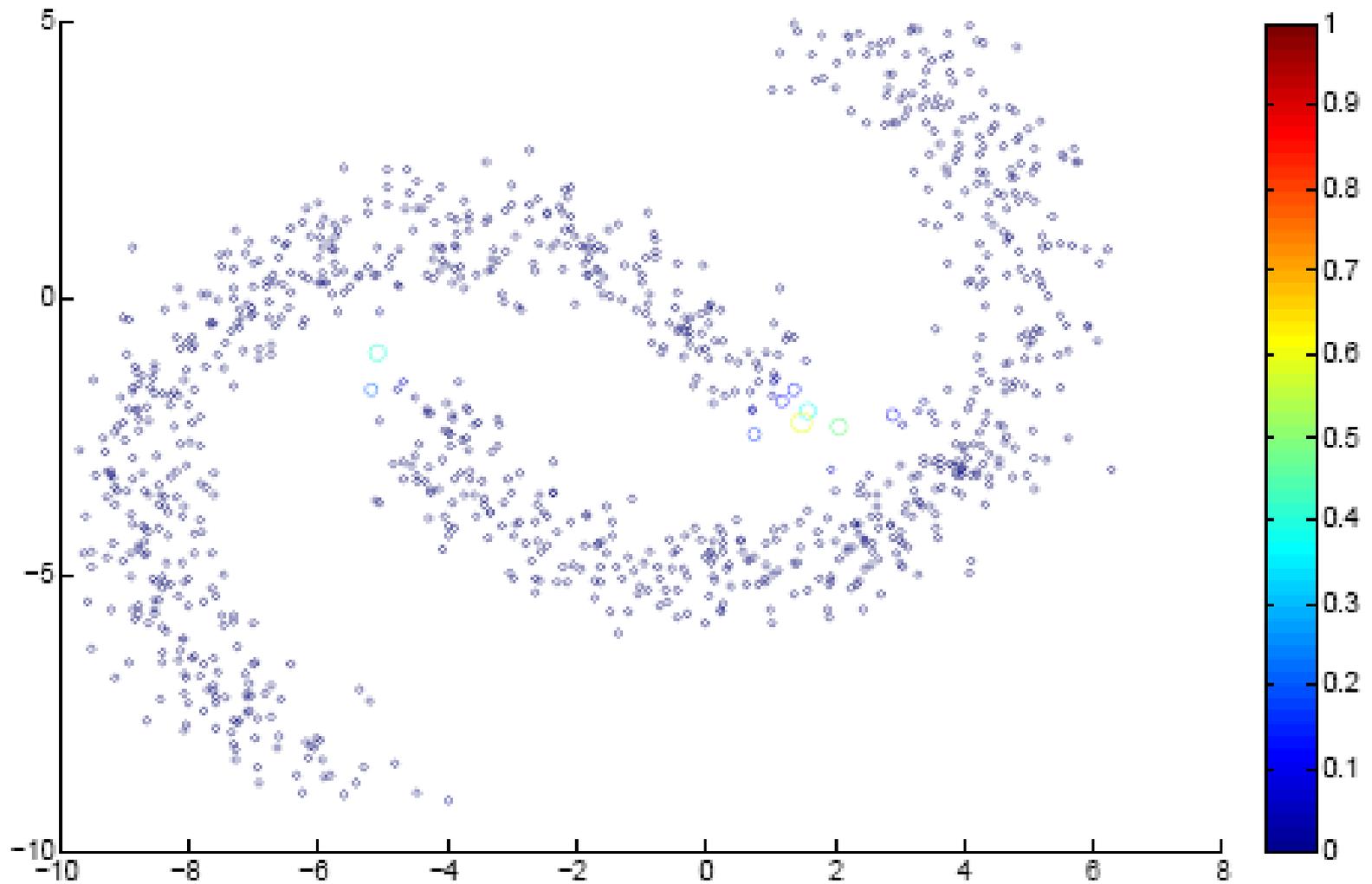
Saída Fuzzy e Detecção de *Outliers*

- Há casos comuns onde alguns nós da rede podem pertencer a mais de uma comunidade
 - Exemplo: Em uma rede social de amizades, indivíduos frequentemente pertencem a várias comunidades: suas famílias, seus colegas de trabalho, seus colegas de escola, etc.
 - Estes nós são chamados *sobrepostos*
 - A maioria dos algoritmos de detecção de comunidade não consegue detectá-los

Saída Fuzzy e Detecção de Outliers

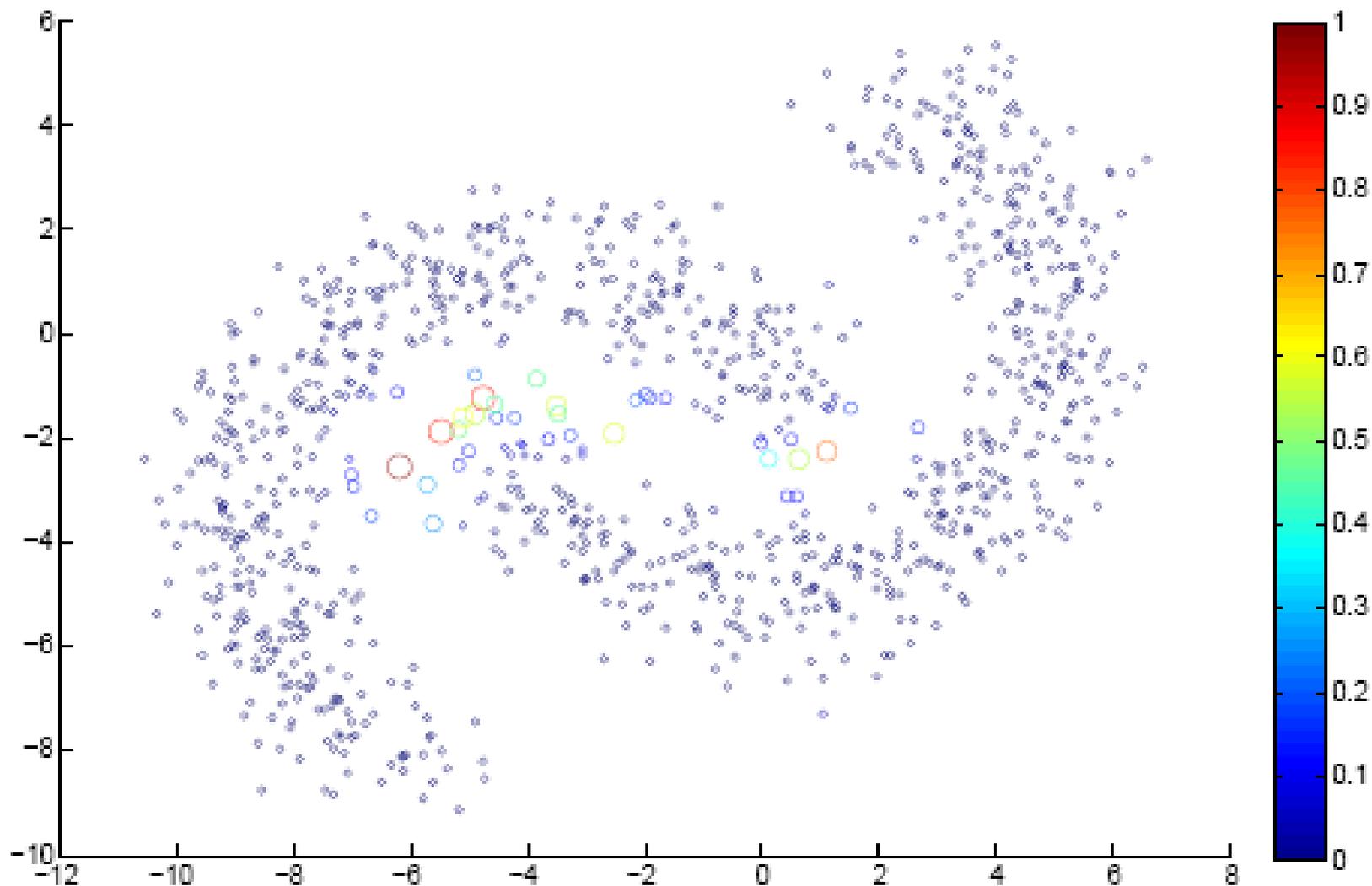
- Algoritmo de partículas padrão
 - Níveis de domínio final definem os rótulos
 - Bastante volátil em certas condições
 - Em nós sobrepostos o time dominante muda frequentemente
 - Níveis não correspondem à medida de sobreposição
- Algoritmo modificado
 - Nova variável: média dos níveis de domínio em cada nó durante todo o tempo
 - Ponderado pela força da partícula
 - Considera apenas o movimento aleatório
 - *Agora o campeão não é mais o time que ganhou os últimos jogos, mas sim o time que ganhou mais jogos durante todo o campeonato*

$$v_i^{\lambda_\ell}(t+1) = v_i^{\lambda_\ell}(t) + \rho_j^\omega(t) \quad f_i^\ell = \frac{v_i^{\lambda_\ell}(\infty)}{\sum_{q=1}^c v_i^{\lambda_q}(\infty)} \quad o_i = \frac{f_i^{\ell^{**}}}{f_i^{\ell^*}}$$



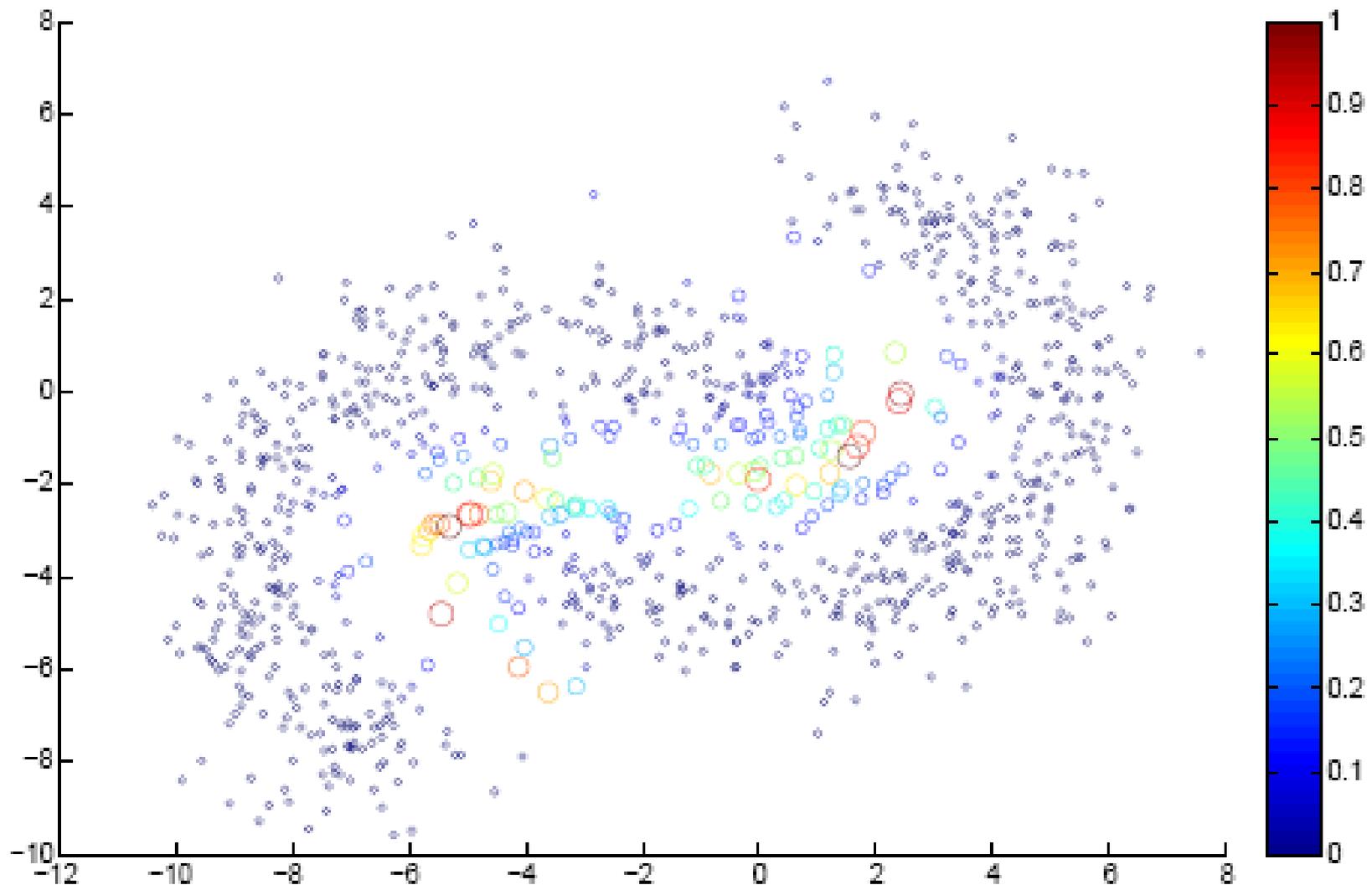
(a)

Classificação fuzzy de classes em forma de banana geradas com diferentes parâmetros de variância: (a) $s = 0.6$ (b) $s = 0.8$ (c) $s = 1.0$. Tamanho e cor dos nós representam seu respectivo índice de sobreposição.



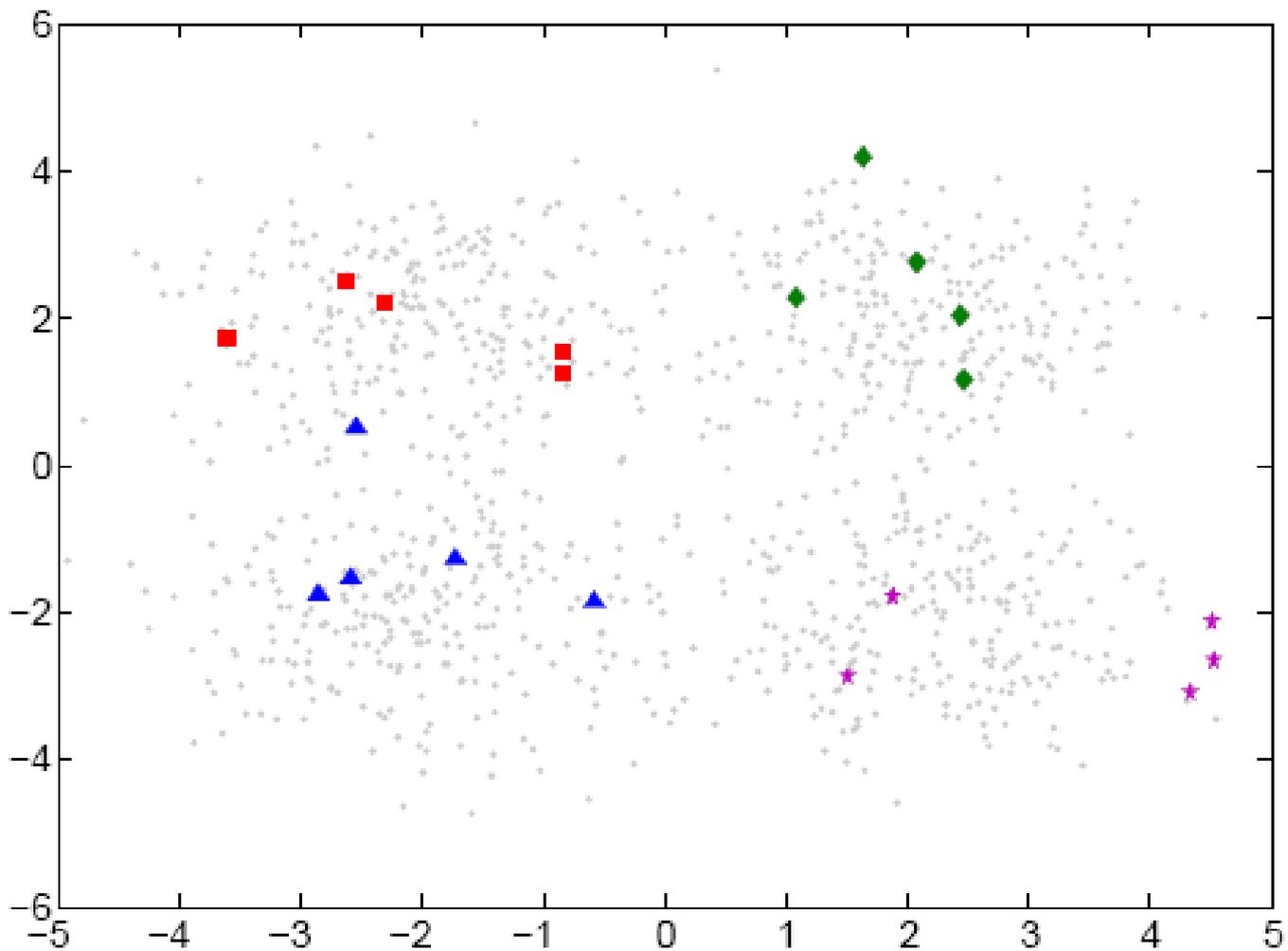
(b)

Classificação fuzzy de classes em forma de banana geradas com diferentes parâmetros de variância: (a) $s = 0.6$ (b) $s = 0.8$ (c) $s = 1.0$. Tamanho e cor dos nós representam seu respectivo índice de sobreposição.



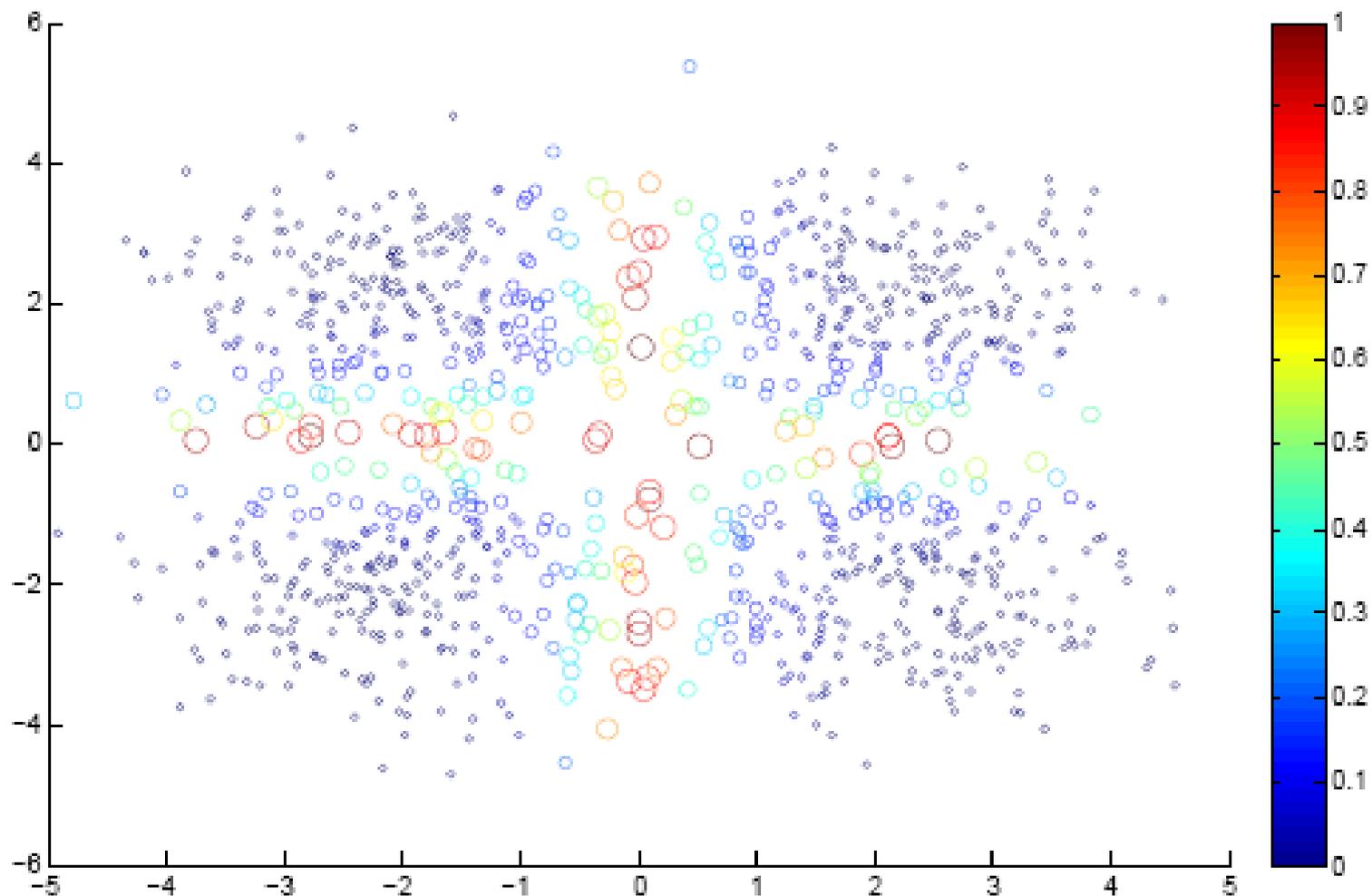
(c)

Classificação fuzzy de classes em forma de banana geradas com diferentes parâmetros de variância: (a) $s = 0.6$ (b) $s = 0.8$ (c) $s = 1.0$. Tamanho e cor dos nós representam seu respectivo índice de sobreposição.



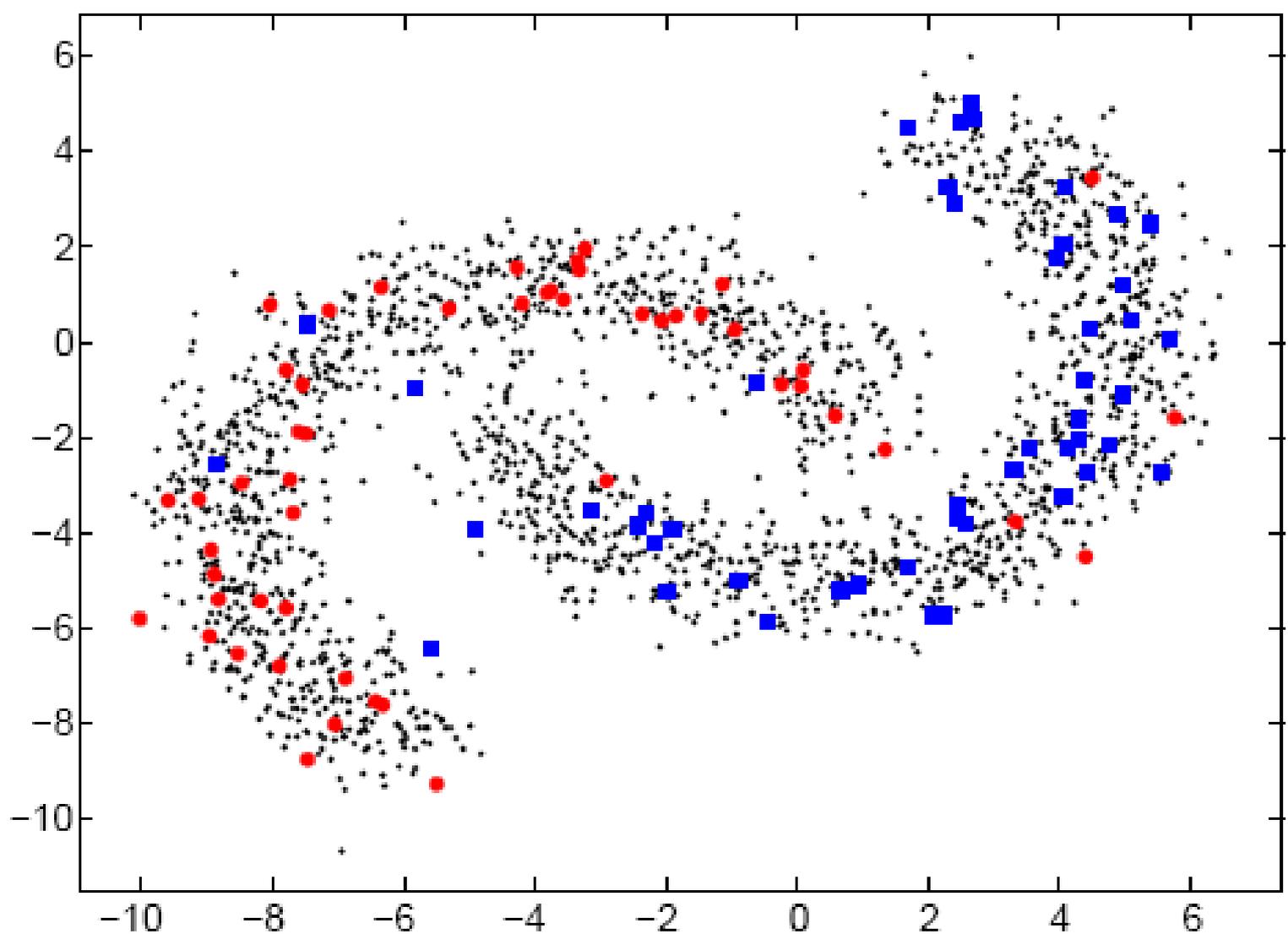
(a)

Classificação de classes com distribuição normal (Gaussiana). (a) base de dados com 1.000 amostras divididas em quatro classes, 20 amostras são rotuladas, 5 de cada classe (quadrados vermelhos, triângulos azuis, losangos verdes e estrelas roxas). (b) tamanho e cores dos nós representam o grau de sobreposição.



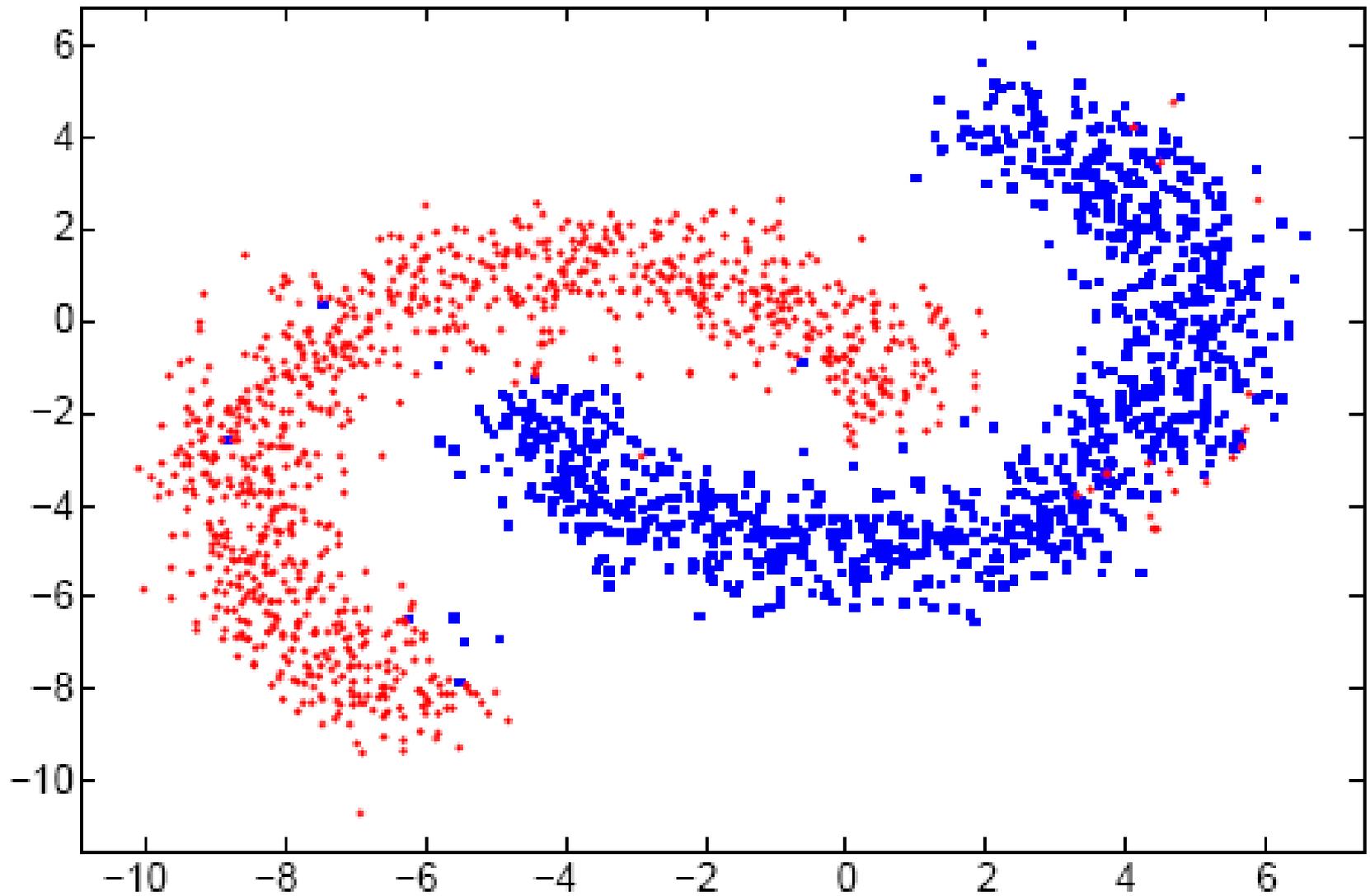
(b)

Classificação de classes com distribuição normal (Gaussiana). (a) base de dados com 1.000 amostras divididas em quatro classes, 20 amostras são rotuladas, 5 de cada classe (quadrados vermelhos, triângulos azuis, losangos verdes e estrelas roxas). (b) tamanho e cores dos nós representam o grau de sobreposição.



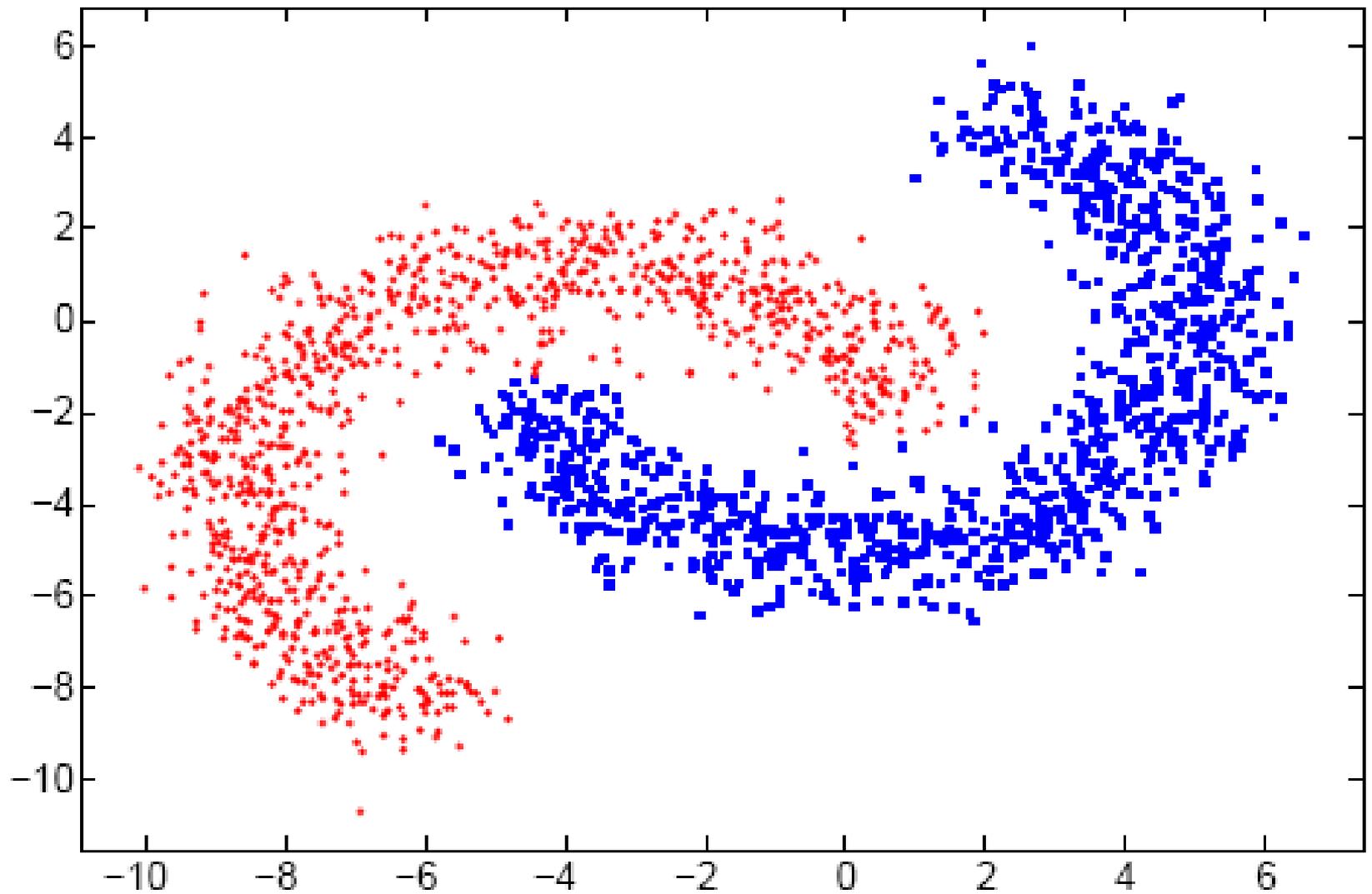
(a)

Comparativo entre o modelo padrão e o modificado: (a) conjunto de dados artificial com alguns nós com rótulo errado (b) classificação pelo método de partículas padrão (c) classificação pelo modelo de partículas modificado



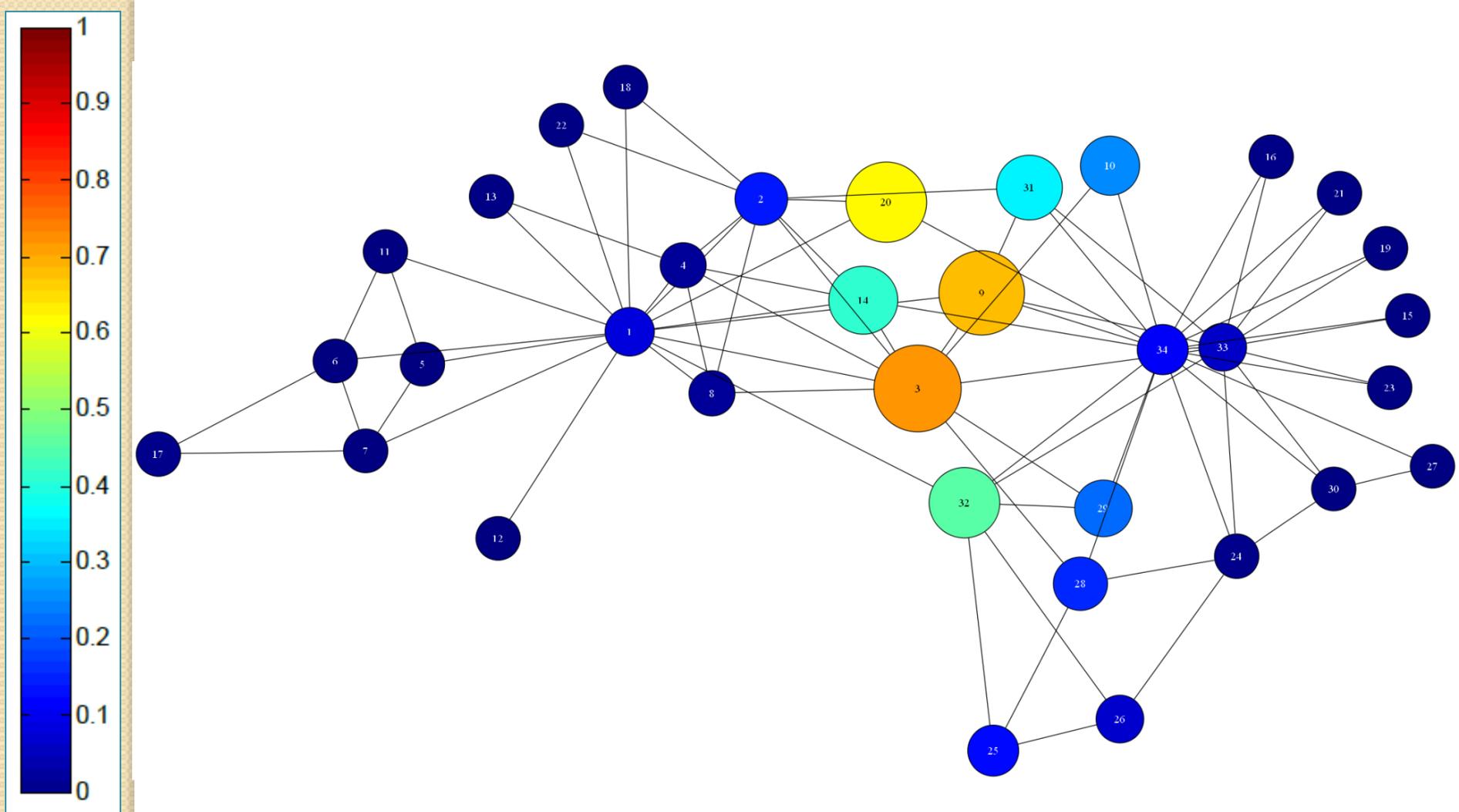
(b)

Comparativo entre o modelo padrão e o modificado: (a) conjunto de dados artificial com alguns nós com rótulo errado (b) classificação pelo método de partículas padrão (c) classificação pelo modelo de partículas modificado



(c)

Comparativo entre o modelo padrão e o modificado: (a) conjunto de dados artificial com alguns nós com rótulo errado (b) classificação pelo método de partículas padrão (c) classificação pelo modelo de partículas modificado



Rede do clube de caratê*. Tamanho e cores dos nós representam seus respectivos índices de sobreposição detectados pelo método de partículas.

* W. W. Zachary, "An information flow model for conflict and fission in small groups," Journal of Anthropological Research, vol. 33, pp. 452–473, 1977.