

В.Д. Козлов, А.И. Майсурадзе

ОБУЧЕНИЕ МЕТРИКИ В ЗАДАЧЕ СОПОСТАВЛЕНИЯ ГРАФОВ

Введение

В современных задачах интеллектуального анализа данных часто возникают ситуации, когда объекты распознавания представляют собой множество компонентов. Одной из важнейших задач для таких объектов является поиск соответствия между отдельными компонентами. При этом, если внутри объектов между компонентами также есть связи, то естественным способом моделирования таких объектов становятся графы, структура которых зависит от конкретной предметной области. При этом отдельные компоненты объекта соответствуют вершинам графа, а отношения между ними — его ребрам. Применительно к графикам задача поиска соответствия между вершинами пары графов называется задачей сопоставления графов.

Графовое представление объектов применяется в задачах анализа и распознавания изображений [1], биометрической идентификации по отпечатку пальца [2] и сетчатке глаза [3], обнаружения вредоносного программного обеспечения [4], молекулярной биологии [5] и др.

В настоящей работе рассматривается задача поиска соответствия между отдельными элементами сложносоставленных объектов. При этом допускается возможность соответствия «один или многие ко многим». Мы предлагаем:

1. новый способ представления информации о предметной области, основанный на графике элементов, что превращает задачу сопоставления компонентов в задачу сопоставления графов;
2. новую модель для решения задачи сопоставления графов, основанную на обучении метрики на вершинах графов, и способ её обучения.

Мы также экспериментально демонстрируем, что предложенная нами модель даёт более высокое качество сопоставления, чем традиционные методы сопоставления графов.

Работа организована следующим образом: проводится обзор существующих методов сопоставления графов, описывается предлагаемый нами новый способ представления информации о

предметной области, предлагается модель для работы с такой информацией, проводится экспериментальное сравнение предлагаемого метода с уже существующими на реальных данных. Завершают работу «Заключение» и список литературы.

Обзор литературы

В простейшем случае *ориентированный граф* представляет собой пару $\langle V, E \rangle$, где V – конечное множество вершин, $E \subseteq V \times V$ – множество рёбер. Расширением этого понятия является так называемый *атрибутированный граф* [6] – это четвёрка $\langle V, E, \mu, \varepsilon \rangle$, где $\mu = \{\mu_1, \dots, \mu_{|V|}\}$ – атрибуты вершин, $\varepsilon = \{\varepsilon_1, \dots, \varepsilon_{|E|}\}$ – атрибуты рёбер. В настоящей работе предполагается, что атрибуты – числовые векторы фиксированной размерности, причем формат и интерпретация атрибутов не меняется от вершины к вершине или от ребра к ребру. Иными словами, для вершин и рёбер задано признаковое описание в формате числовых векторов. Задача сопоставления графов G_1 и G_2 состоит в поиске отображения $h : V_1 \rightarrow V_2$.

Традиционно задача сопоставления графов формулируется как задача поиска соответствия между вершинами, минимизирующего некоторый функционал качества [7]. Как правило, предполагается, что для вершин $i \in V_1$ и $i' \in V_2$ определена стоимость сопоставления $c_{ii'}$, а для пар вершин $(i, i'), (j, j') \in V_1 \times V_2$ – стоимость сопоставления $d_{ii'jj'}$. При этом задача сопоставления эквивалентна квадратичной задаче о назначениях [8]: $\sum_{i,i'} c_{ii'}x_{i,i'} + \sum_{(i,i'),(j,j')} d_{ii'jj'}x_{i,i'}x_{j,j'} \rightarrow \max_x$, где x – бинарная матрица размера $|V_1| \times |V_2|$. В литературе были предложены различные способы приближённого решения такой задачи с использованием спектральных свойств матрицы стоимости [9, 10], градиентного спуска с учётом особенностей задачи [11, 12], метода чередующихся проекций [13] и метода отжига [1].

Методы машинного обучения в задаче сопоставления графов до настоящего времени применялись довольно ограниченно. В основном они использовались для получения стоимости сопоставлений $c_{ii'}$ и $d_{ii'jj'}$ [14–16].

Универсальный граф

В настоящей работе рассматривается задача поиска соответствия между отдельными компонентами сложного объекта. Мы предлагаем новый способ обогащения задачи информацией о предметной области.

Будем считать, что имеется некоторое конечное и заранее фиксированное множество типов компонентов. Тогда предлагается ввести связи между типами компонентов в соответствии с информацией о предметной области и требованиями задачи. Полученная структура представляет собой граф, называемый в дальнейшем *универсальным*

графом. Тогда каждый конкретный объект можно представить в виде графа, являющегося подграфом универсального графа: компоненты объекта соединены ребром, если типы компонентов в универсальном графе соединены ребром. Пример универсального графа и графа конкретного объекта представлен на рис. 1.

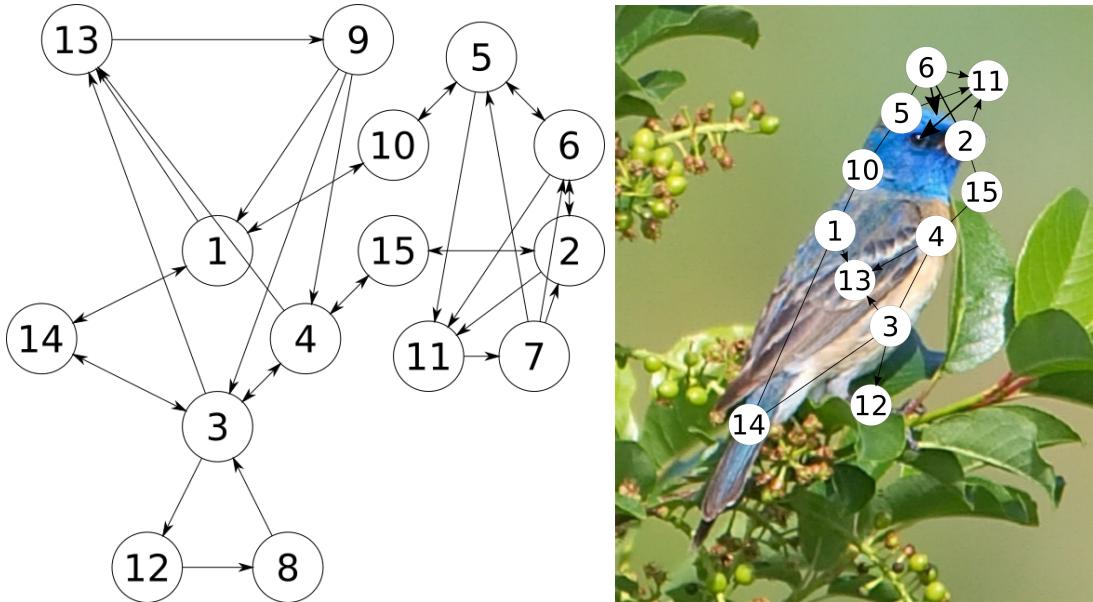


Рисунок 1. Универсальный граф (слева) и граф объекта (справа). Объектами являются изображения птиц, на которых выделены различные части тела. Имеется 15 различных типов компонентов (части тела птицы), причём некоторые части тела на конкретной фотографии могут не появляться.

Таким образом, объекты с учётом информации из предметной области представляют собой граф, и задача поиска соответствия между ними может быть сформулирована как задача сопоставления графов. Ниже будет предложен метод машинного обучения для решения задачи сопоставления графов.

Модель сопоставления графов

Для сопоставления графов объектов мы предлагаем модель на основе обучения метрики между представлениями вершин графов. Модель можно схематически представить в следующем виде:

1. получение первичного представления компонентов без использования информации о графе;
2. получение вторичного представления компонентов на основе информации о графе;
3. вычисление сходства полученных представлений вершин, их сопоставление.

Заметим, что наша схема сопоставляет вершины исключительно на основе их вторичного представления и не сводит задачу к квадратичной задаче о назначениях.

Модель получения первичного представления компонентов зависит от конкретной задачи. Поэтому подход, применявшийся в процессе экспериментов в настоящей работе, будет описан в разделе «Описание экспериментов». Результатом работы этой части модели является набор векторов $\{\mu_1, \dots, \mu_{|V|}\}$ — признаковых описаний компонентов. Также на этом этапе возможно получение признаковых описаний для пар компонентов, из которого в дальнейшем можно получить описания рёбер $\{\varepsilon_1, \dots, \varepsilon_{|E|}\}$.

Для получения вторичного представления мы предлагаем использовать графовые свёрточные сети [17]. Один слой такой сети в общем случае принимает граф $G = \langle V, E \rangle$ и пересчитывает атрибуты вершин $\{\mu_1, \dots, \mu_{|V|}\}$ и рёбер $\{\varepsilon_1, \dots, \varepsilon_{|E|}\}$ согласно формулам:

$$\begin{aligned} \mu_k^{(n+1)} &= \sigma \left(W_{vself}^{(n)} \mu_k^{(n)} + \frac{1}{|N_{in}(k)|} \sum_{i \in N_{in}(k)} \left(W_{vin}^{(n)} \mu_i^{(n)} + W_{ein}^{(n)} \varepsilon_{i \rightarrow k}^{(n)} \right) + \right. \\ &\quad \left. \frac{1}{|N_{out}(k)|} \sum_{i \in N_{out}(k)} \left(W_{vout}^{(n)} \mu_i^{(n)} + W_{eout}^{(n)} \varepsilon_{k \rightarrow i}^{(n)} \right) + b_v^{(n)} \right) \quad (1) \end{aligned}$$

$$\varepsilon_{i \rightarrow j}^{(n+1)} = \sigma \left(W_{eself}^{(n)} \varepsilon_{i \rightarrow j}^{(n)} + W_{vfrom}^{(n)} \mu_i^{(n)} + W_{vto}^{(n)} \mu_j^{(n)} + b_e^{(n)} \right) \quad (2)$$

где $N_{in}(k)$, $N_{out}(k)$ — вершины, из которых рёбра идут в вершину k и в которые выходят из вершины k соответственно. Заметим, что если у вершины соседей нет, то слой эквивалентен обычному полносвязному. Результатом работы этой части модели является набор векторов $\{\mu'_1, \dots, \mu'_{|V|}\}$ — признаковых описаний компонентов, построенных с учётом графовой структуры объекта.

Для получения сопоставлений между вершинами графов G_1 и G_2 мы предлагаем вычислять попарные евклидовые расстояния между векторами вторичных описаний компонентов $\mu_i'^1$ и $\mu_{i'}'^2$. Если расстояние ниже определённого порога, то предполагается, что соответствующие вершины сопоставляются. Как правило, в задачах обучения метрики рассматривается метрика Махalanобиса [18]; в нашем случае обучение метрики на вершинах де-факто происходит на этапе вторичного представления. Результатом работы этой части модели является вещественная матрица D размера $|V_1| \times |V_2|$ попарных расстояний между вершинами графов, а также бинарная матрица сопоставлений M того же размера, получаемая описанным выше методом дискретизации матрицы D .

Обучающая выборка состоит из пар объектов, для которых известны графы G_1 и G_2 , основанные на универсальном графе, а также

бинарная матрица действительных сопоставлений H . В качестве функции потерь для обучения модели мы предлагаем $L(H, D) = \|H - \exp(-D^2)\|_2^2 / (|V_1| * |V_2|)$. Структура решения позволяет использовать метод обратного распространения ошибки для обучения вторичного представления и, де-факто, метрики на вершинах.

Отметим, что такой подход принципиально отличается от традиционного, основанного на решении квадратичной задачи о назначениях. Мы отказываемся на этапе сопоставления от информации о связях и используем её только для обучения представления вершин.

Описание задачи

Мы демонстрируем работу нашего метода на примере задачи сопоставления точек на изображениях. Для этого мы используем набор изображений CUB_200_2011 [19]. Набор состоит из фотографий птиц различных видов. На каждой фотографии отмечено не более 15 точек – различных частей тела птицы, для каждой точки известен её тип. При этом каждая часть тела может либо присутствовать на фотографии 1 раз, либо не присутствовать вовсе.

В данной задаче мы считаем, что компоненты соответствуют друг другу, если части тела имеют одинаковые функции. Схожесть функций частей тела определяется на основе анализа информации о виде птицы. Например, мы считаем, что лапы водоплавающей птицы не должны сопоставляться с лапами певчей, или что клюв хищной птицы не должен сопоставляться с клювом насекомоядной. При этом мы конструируем универсальный граф на типах частей тела так, чтобы в нём содержалась информация о строении тела птицы. Предложенный универсальный граф и порождённый им график конкретного изображения представлены на рис. 1. Таким образом, графы в задаче строятся на основе морфологической информации, но требуется обучить модель сопоставлению на основе функций частей тела.

Описание экспериментов

Для получения первичных представлений компонентов мы используем стандартную предварительно обученную модель для анализа изображений MobileNetV2 [20]. С её помощью мы преобразуем исходное изображение в карту признаков. Далее элементы этой карты в координатах, соответствующих координатам частей на исходном изображении, используются как первичные признаковые описания компонентов. Во всех экспериментах веса модели для получения первичных описаний фиксируются, если не сказано обратное.

Мы рассматриваем следующие вариации моделей:

1. модель с использованием только информации о наличии рёбер: у модели отсутствуют атрибуты рёбер, в формуле пересчёта атрибутов

- вершин (1) не используются атрибуты рёбер ε ;
2. модель с использованием атрибутов рёбер: каждое ребро описывается атрибутом – евклидовой длиной (расстояние между соответствующими точками на изображении);
 3. модель с дообучением первичного представления: использование глубокого обучения позволяет настраивать веса первой части модели совместно с последующими.

Каждая последующая модель является расширением предыдущей.

В качестве базовой модели мы используем адаптацию модели сопоставления, предложенной в работе [16], к нашей задаче. Эта модель основана на обучении коэффициентов стоимости сопоставлений $c_{ii'}$ и $d_{ii'jj'}$ с последующим решением квадратичной задачи о назначениях. Для получения коэффициентов стоимости мы используем уже описанный механизм графовой сети. После получения вторичных представлений вершин $\{\mu^{1'}_i\}, \{\mu^{2'}_{i'}\}$ и рёбер $\{\varepsilon^{1'}_{(i,j)}\}, \{\varepsilon^{2'}_{(i',j')}\}$ коэффициентов стоимости рассчитываются как $c_{ii'} = \frac{\langle \mu^{1'}_i, \mu^{2'}_{i'} \rangle}{\|\mu^{1'}_i\|_2 \|\mu^{2'}_{i'}\|_2}$ (косинусное расстояние) и $d_{ii'jj'} = \exp(-\|\varepsilon^{1'}_{(i,j)} - \varepsilon^{2'}_{(i',j')}\|_2^2)$. В такой модели представления для вершин и рёбер, использующиеся для получения коэффициентов стоимости, получаются тем же способом, что и в предлагаемых нами моделях, что позволяет сравнивать качество сопоставления.

Мы использовали два показателя качества сопоставления:

- общая точность сопоставления:

$$Acc(H, M) = \frac{1}{|V_1||V_2|} \sum_{i \in V_1, i' \in V_2} [H_{ii'} = M_{ii'}];$$
- мера Жаккара: $Jacc(H, M) = \frac{|H \cap M|}{|H \cup M|}$.

Здесь бинарные матрицы сопоставления H (действительные сопоставления) и M (предсказанные сопоставления) также рассматриваются как некоторые подмножества всех возможных сопоставлений $V_1 \times V_2$. Применение меры Жаккара обусловлено тем, что мощность множества верных сопоставлений H невелика по сравнению с мощностью множества $V_1 \times V_2$ (в матрице H много нулей).

В процессе экспериментов каждая модель обучалась на случайных парах изображений из обучающей части набора данных. Качество сопоставления рассчитывалось на случайных парах изображений из тестовой части набора данных. Таким образом, изображения, использовавшиеся при обучении, не встречались при тестировании, и наоборот. Поскольку оба графа объектов структурно являются подграфами универсального графа, естественно ожидать, что качество сопоставления тем выше, чем выше вершин совпадает. Поэтому мы будем

исследовать качество сопоставления в зависимости от числа общих вершин.

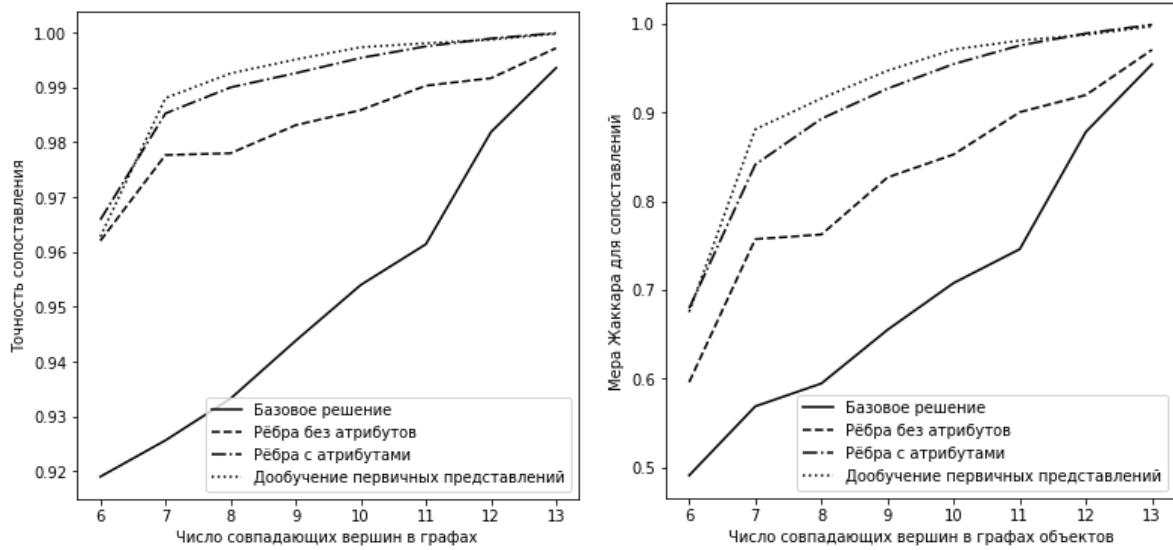


Рисунок 2. Общая точность и мера Жаккара в зависимости от числа совпадающих вершин в графах.

Результаты

На графиках на рис. 2 приведены показатели качества в зависимости от числа совпадающих вершин в графах.

Хорошо видно, что предложенный нами метод на основе обучения метрики на вершинах графов даже в ослабленном варианте (без использования атрибутов рёбер) показывает более высокое качество, чем метод, основанный на решении квадратичной задачи о назначениях. Кроме того, видно, что добавление информации об атрибутах рёбер позволяет заметно улучшить качество модели по сравнению с моделью, не использующей эту информацию: даже при всего 7 совпадающих вершинах (половина числа вершин в универсальном графе) модель показывает точность 0.98 и меру Жаккара 0.84, то есть модель устойчива к эрозии. При этом добавление возможности совместного обучения модели получения первичного представления компонентов незначительно влияет на качество сопоставления.

Таким образом, можно сделать вывод, что предложенный метод сопоставления даёт лучшие результаты, чем традиционный.

Заключение

В работе предложен новый метод обогащения задачи информацией о предметной области на основе графов. Предложена новая модель работы с графовыми данными, основанная на обучении метрики на вершинах. Продемонстрировано, что такая модель даёт лучшие результаты, чем традиционные модели сопоставления графов, и обладает высокой устойчивостью к вариативности в данных.

Литература

1. Zhou F., de la Torre F. Factorized graph matching. // 2012 IEEE Conference on Computer Vision and Pattern Recognition. — IEEE, 2012. — P. 127–134.
2. Neuhaus M., Bunke H. A graph matching based approach to fingerprint classification using directional variance. // Lecture Notes in Computer Science. — Vol. 3546. — Springer-Verlag Berlin, 2005. — P. 191–200.
3. Lajevardi S., Arakala A., Davis S., Horadam K. Retina Verification System Based on Biometric Graph Matching. // IEEE Transactions on Image Processing. — Vol. 22. — IEEE, 2013. — P. 3625–3635.
4. Bruschi D., Martignoni L., Monga M. Detecting self-mutating malware using control-flow graph matching. // Lecture Notes in Computer Science. — Vol. 4064. — Springer-Verlag Berlin, 2006. — 129–143
5. Yan X., Yu P. S., Han J. Substructure similarity search in graph databases. // Proceedings of the ACM SIGMOD International Conference on Management of Data. — ACM Press, 2005. — P. 766–777.
6. Tsai W.-H., Fu K.-S. Error-correcting isomorphisms of attributed relational graphs for pattern analysis. // IEEE Transactions on systems, man, and cybernetics. — Vol. 9, No. 12. — IEEE, 1979. — P. 757–768.
7. Yan J., Yin X.-Ch., Lin W., Deng Ch., Zha H., Yang X. A Short Survey of Recent Advances in Graph Matching. // Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval. — ACM, 2016. — P. 167–174.
8. Lawler E. L. The Quadratic Assignment Problem. // Management Science. — Vol. 9, No. 4. — 1963. — P. 586–599.
9. Leordeanu M., Hebert M. A spectral technique for correspondence problems using pairwise constraints. // Proceedings of the IEEE International Conference on Computer Vision. — Vol. II. — IEEE, 2005. — P. 1482–1489.
10. Cour T., Srinivasan P., Shi J. Balanced Graph Matching. // Advances in Neural Information Processing Systems. — MIT Press, 2006. — P. 313–320.
11. Gold S., Rangarajan A. A graduated assignment algorithm for graph matching. // IEEE Transactions on Pattern Analysis and Machine Intelligence. — Vol. 18, No. 4. — IEEE, 1996. — P. 377–388.
12. Leordeanu M., Hebert M., Sukthankar R. An Integer Projected Fixed Point Method for Graph Matching and MAP Inference. // Advances in Neural Information Processing Systems. — Vol. 1, No. 3. — MIT Press, 2009. — P. 1114–1122.

13. *van Wyk B. J., van Wyk M. A.* A POCS-based graph matching algorithm. // IEEE Transactions on Pattern Analysis and Machine Intelligence. — Vol. 26, No. 11. — IEEE, 2004. — P. 1526–1530.
14. *Caetano T. S., McAuley J. J., Cheng L., Le Q. V., Smola A. J.* Learning graph matching. // IEEE Transactions on Pattern Analysis and Machine Intelligence. — Vol. 31, No. 6. — IEEE, 2009. — P. 1048–1058.
15. *Leordeanu M., Sukthankar R., Hebert M.* Unsupervised learning for graph matching. // International Journal of Computer Vision. — Vol. 96, No. 1. — Springer Science+Business Media, LLC, 2012. — P. 28–45.
16. *Zanfir A., Sminchisescu C.* Deep Learning of Graph Matching. // The IEEE Conference on Computer Vision and Pattern Recognition. — IEEE, 2018. — P. 2684–2693.
17. *Kipf T. N., Welling M.* Semi-supervised classification with graph convolutional networks. // 5th International Conference on Learning Representations, ICLR 2017 – Conference Track Proceedings. — 2019.
18. *Kulis B.* Metric Learning: A Survey. // Foundations and Trends in Machine Learning. — Vol. 5, No. 4. — Now Publishers Inc, 2013. — P. 287–364.
19. *Wah C., Branson S., Welinder P., Perona P., Belongie S.* The Caltech-UCSD Birds-200-2011 Dataset. // California Institute of Technology, 2011.
20. *Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L. Ch.* MobileNetV2: Inverted Residuals and Linear Bottlenecks. // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. — IEEE, 2018. — P. 4510–4520.