Publications of Nicolas Klodt

This document lists all peer-reviewed publications of Nicolas Klodt, Chair for Algorithm Engineering, Hasso Plattner Institute, Potsdam, Germany. This listing was automatically generated on October 10, 2024. An up-to-date version is available online at hpi.de/friedrich/docs/publist/klodt.pdf.

Journal articles

[1] Friedrich, T., Göbel, A., Klodt, N., Krejca, M. S., Pappik, M., Analysis of the survival time of the SIRS process via expansion. In: *Electronic Journal of Probability* 29, pp. 1–29, 2024.

We study the SIRS process—a continuous-time Markov chain modeling the spread of infections on graphs. In this model, vertices are either susceptible, infected, or recovered. Each infected vertex becomes recovered at rate 1 and infects each of its susceptible neighbors independently at rate λ, and each recovered vertex becomes susceptible at a rate ϱ, which we assume to be independent of the graph size. A central quantity of the SIRS process is the time until no vertex is infected, known as the survival time. Surprisingly though, to the best of our knowledge, all known rigorous theoretical results that exist so far immediately carry over from the related SIS model and do not completely explain the behavior of the SIRS process. We address this imbalance by conducting theoretical analyses of the SIRS process via the expansion properties of the underlying graph. Our first result shows that the expected survival time of the SIRS process on stars is at most polynomial in the graph size for any value of λ. This behavior is fundamentally different from the SIS process, where the expected survival time is exponential already for small infection rates. This raises the question of which graph properties result in an exponential survival time. Our main result is an exponential lower bound of the expected survival time of the SIRS process on expander graphs. Specifically, we show that on expander graphs G with n vertices, degree close to d, and sufficiently small spectral expansion, the SIRS process has expected survival time at least exponential in n when $\lambda \geq c/d$ for a constant $c > 1$. Previous results on the SIS process show that this bound is almost tight. Additionally, our result holds even if G is a subgraph. Notably, our result implies an almost-tight threshold for Erdos-Renyi-graphs and a regime of exponential survival time for complex network models. The proof of our main result draws inspiration from Lyapunov functions used in mean-field theory to devise a two-dimensional potential function and from applying a negative-drift theorem to show that the expected survival time is exponential.

Conference papers

[2] Friedrich, T., Göbel, A., Klodt, N., Krejca, M. S., Pappik, M., The Irrelevance of Influencers: Information Diffusion with Re-Activation and Immunity Lasts Exponentially Long on Social Network Models. In: *Annual AAAI Conference on Artificial Intelligence*, 2024.

Information diffusion models on networks are at the forefront of AI research. The dynamics of such models typically follow stochastic models from epidemiology, used to model not only infections but various phenomena, including the behavior of computer viruses and viral marketing campaigns. A core question in this setting is how to efficiently detect the most influential vertices in the host graph such that the infection survives the longest. In processes that incorporate re-infection of the vertices, such as the SIS process, theoretical studies identify parameter thresholds where the survival time of the process rapidly transitions from logarithmic to super-polynomial. These results contradict the intuition that the starting configuration is relevant, since the process will always either die out fast or survive almost indefinitely. A shortcoming of these results is that models incorporating short-term immunity (or creative advertisement fatigue) have not been subjected to such a theoretical analysis so far. We reduce this gap in the literature by studying the SIRS process, a more realistic model, which besides re-infection additionally incorporates short-term immunity. On complex network models, we identify parameter regimes for which the process survives exponentially long, and we get a tight threshold for random graphs. Underlying these results is our main technical contribution, showing a threshold behavior for the survival time of the SIRS process on graphs with large expander subgraphs, such as social network models.

[3] Angrick, S., Bals, B., Friedrich, T., Gawendowicz, H., Hastrich, N., Klodt, N., Lenzner, P., Schmidt, J., Skretas, G., Wells, A., How to Reduce Temporal Cliques to Find Sparse Spanners. In: *European Symposium on Algorithms (ESA)*, 2024.

[4] Friedrich, T., Göbel, A., Klodt, N., Krejca, M. S., Pappik, M., From Market Saturation to Social Reinforcement: Understanding the Impact of Non-Linearity in Information Diffusion Models. In: *The 23rd International Conference on Autonomous Agents and Multi-Agent Systems*, 2024.

Diffusion of information in networks is at the core of many problems in AI. Common examples include the spread of ideas and rumors as well as marketing campaigns. Typically, information diffuses at a non-linear rate, for example, if markets become saturated or if users of social networks reinforce each other's opinions. Despite these characteristics, this area has seen little research, compared to the vast amount of results for linear models, which exhibit less complex dynamics. Especially, when considering the possibility of re-infection, no fully rigorous guarantees exist so far. We address this shortcoming by studying a very general non-linear diffusion model that captures saturation as well as reinforcement. More precisely, we consider a variant of the SIS model in which vertices get infected at a rate that scales polynomially in the number of their infected neighbors, weighted by an infection coefficient λ. We give the first fully rigorous results for thresholds of λ at which the expected survival time becomes super-polynomial. For cliques we show that when the infection rate scales sub-linearly, the threshold only shifts by a poly-logarithmic factor, compared to the standard SIS model. In contrast, super-linear scaling changes the process considerably and shifts the threshold by a polynomial term. For stars, sub-linear and super-linear scaling behave similar and both shift the threshold by a polynomial factor. Our bounds are almost tight, as they are only apart by at most a poly-logarithmic factor from the lower thresholds, at which the expected survival time is logarithmic.

[5] Bilò, D., Cohen, S., Friedrich, T., Gawendowicz, H., Klodt, N., Lenzner, P., Skretas, G., Temporal Network Creation Games. In: *International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 2511–2519, 2023.

Most networks are not static objects, but instead they change over time. This observation has sparked rigorous research on temporal graphs within the last years. In temporal graphs, we have a fixed set of nodes and the connections between them are only available at certain time steps. This gives rise to a plethora of algorithmic problems on such graphs, most prominently the problem of finding temporal spanners, i.e., the computation of subgraphs that guarantee all pairs reachability via temporal paths. To the best of our knowledge, only centralized approaches for the solution of this problem are known. However, many real-world networks are not shaped by a central designer but instead they emerge and evolve by the interaction of many strategic agents. This observation is the driving force of the recent intensive research on game-theoretic network formation models. In this work we bring together these two recent research directions: temporal graphs and game-theoretic network formation. As a first step into this new realm, we focus on a simplified setting where a complete temporal host graph is given and the agents, corresponding to its nodes, selfishly create incident edges to ensure that they can reach all other nodes via temporal paths in the created network. This yields temporal spanners as equilibria of our game. We prove results on the convergence to and the existence of equilibrium networks, on the complexity of finding best agent strategies, and on the quality of the equilibria. By taking these first important steps, we uncover challenging open problems that call for an in-depth exploration of the creation of temporal graphs by strategic agents.

[6] Berger, J., Böther, M., Doskoč, V., Gadea Harder, J., Klodt, N., Kötzing, T., Lötzsch, W., Peters, J., Schiller, L., Seifert, L., Wells, A., Wietheger, S., Learning Languages with Decidable Hypotheses. In: *Computability in Europe (CiE)*, pp. 25–37, 2021.

In language learning in the limit, the most common type of hypothesis is to give an enumerator for a language, a W-index. These hypotheses have the drawback that even the membership problem is undecidable. In this paper, we use a different system which allows for naming arbitrary decidable languages, namely programs for characteristic functions (called C-indices). These indices have the drawback that it is now not decidable whether a given hypothesis is even a legal C-index. In this first analysis of learning with C-indices, we give a structured account of the learning power of various restrictions employing C-indices, also when compared with W-indices. We establish a hierarchy of learning power depending on whether C-indices are required (a) on all outputs; (b) only on outputs relevant for the class to be learned or (c) only in the limit as final, correct hypotheses. We analyze all these questions also in relation to the mode of data presentation. Finally, we also ask about the relation of semantic versus syntactic convergence and derive the map of pairwise relations for these two kinds of convergence coupled with various forms of data presentation.

[7] Casel, K., Friedrich, T., Issac, D., Klodt, N., Seifert, L., Zahn, A., A Color-blind 3-Approximation for Chromatic Correlation Clustering and Improved Heuristics. In: *Knowledge Discovery and Data Mining (KDD)*, pp. 882–891, 2021.

Chromatic Correlation Clustering (CCC) models clustering of objects with categorical pairwise relationships. The model can be viewed as clustering the vertices of a graph with edge-labels (colors). Bonchi et al. [KDD 2012] introduced it as a natural generalization of the well studied problem Correlation Clustering (CC), motivated by real-world applications from data-mining, social networks and bioinformatics. We give theoretical as well as practical contributions to the study of CCC. Our main theoretical contribution is an alternative analysis of the famous Pivot algorithm for CC. We show that, when simply run color-blind, Pivot is also a linear time 3-approximation for CCC. The previous best theoretical results for CCC were a 4-approximation with a high-degree polynomial runtime and a linear time 11-approximation, both by Anava et al. [WWW 2015]. While this theoretical result justifies Pivot as a baseline comparison for other heuristics, its blunt color-blindness performs poorly in practice. We develop a color-sensitive, practical heuristic we call Greedy Expansion that empirically outperforms all heuristics proposed for CCC so far, both on real-world and synthetic instances. Further, we propose a novel generalization of CCC allowing for multi-labelled edges. We argue that it is more suitable for many of the real-world applications and extend our results to this model.