

Data Driven Assembly of Procedurally Modeled Facilities

M. S. Bishop^{1,2}, J. Ferrer² and N. Max¹

¹University of California, Davis, USA

²Universitat Politècnica de València (UPV), Spain

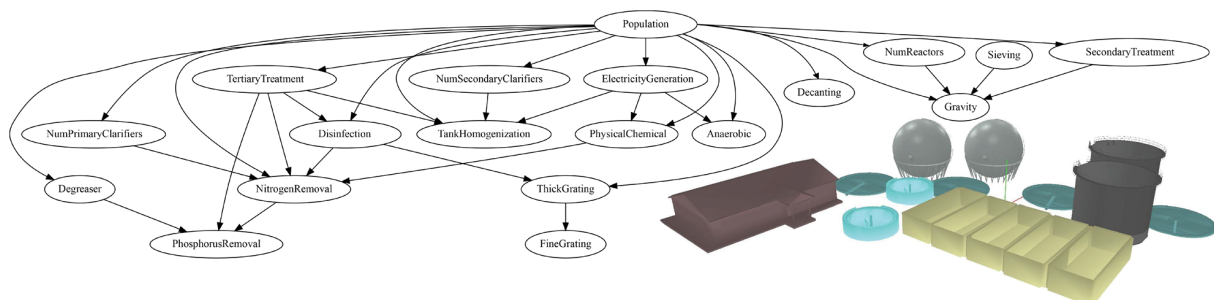


Figure 1: Bayesian network (number of nodes reduced for display) of a facility and a set of arranged tri-mesh models [Tri13].

Abstract

We present a method to arrange components of industrial facilities in a constrained site footprint. We use a probabilistic graphical model of industrial sites and existing procedural modeling methods to automate the assembly and 3D modeling of wastewater treatment plants. A knowledge engineered approach produces a combination of components that inherently contains domain specific information like process dependencies and facility size. The inferred combination is laid out using mathematical optimization or via a physics-based simulation resulting in an arrangement that respects the industrial process and design plausibility.

Categories and Subject Descriptors (according to ACM CCS): I.3.2 [Computer Graphics]: Graphics Systems—Stand-alone Systems, G.3 [Probability and Statistics]: Probabilistic Algorithms—Bayesian Networks, I.2.6 [Computing Methodologies]: Learning—Knowledge Acquisition.

1. Introduction

The built world is not easy to re-model in 3D [VAW*10]. Research methods that try to automate the rebuilding process generally start from images [BM99, FRL*98, PKVG00]; LiDAR data [CL96, FNSZ06]; or use procedural methods to model buildings [MWH*06] and detailed facades [HWA*10, KW11, MZVWG07]. In this work we focus on synthesizing industrial facilities that require teams of experts in disparate fields [BH93] to construct. To address some of the complexity in making 3D facility models we propose a probabilistic approach to component selection, optimized and physics-based layouts, and 3D models with procedurally generated and triangle-mesh geometry. To demonstrate the usefulness of our approach we chose wastewater treatment

plants (WWTP) as the example facility type because design documents are publically available, and the structures (clarifiers, reactors, etc.) are geometrically distinct and visible from aerial photographs. The primary motivation for this work is to use the 3D facility models as input to synthetic aerial image generators [Cen]. The 2D images can then be used as benchmark imagery for verification and validation [GDSS*11] of automated detection, classification, and labeling algorithms [RTP*10, RPV*11].

2. Related Work

Procedural methods to automatically generate 3D building models tend to focus on residential and commercial

buildings [PM01, WMW*08], reconstructing cultural sites [HMG09], creating facades [TKS*13], and enforcing structural engineering constraints [WOD09]. There are probabilistic methods to assemble individual 3D models from segmented and labeled shape libraries [CKGK11, KCKK12], but there are few automated methods to generate plausible industrial 3D facility models. One approach outlined by Nomura and Miyata [NM11] generates 3D models of process plants from silhouette sketches of facility images, but their approach does not respect the facility process functionality and the resulting geometry may contain intersections. In addition to generating geometry, 3D facility modeling considers interconnected components that interact in some fashion, e.g., an assembly line, and while not directly related to facilities, there is interest from the graphics community in the related problem of furniture layout [GS09, MSL*11], open world layouts [YYW*12] and creating new 3D arrangements from existing 3D scenes [FRS*12] that might be adapted to work for facility layout and modeling.

3. Approach

Features of several hundred WWTP's are used to construct a Bayesian network (BN). The use of a BN is inspired by Merrell et al. [MSK10] who use a probabilistic method to derive residential layouts from user specifications. Merrell et al. combine BN structure learning and optimization to compute new residential layouts from a corpus of existing homes. A sampling from our BN provides the set of facility components given some high level user constraints like the population in the target area. Then, a layout algorithm is applied to the set of components, a 2D footprint generated and a 3D model constructed. (See Figure 2.)

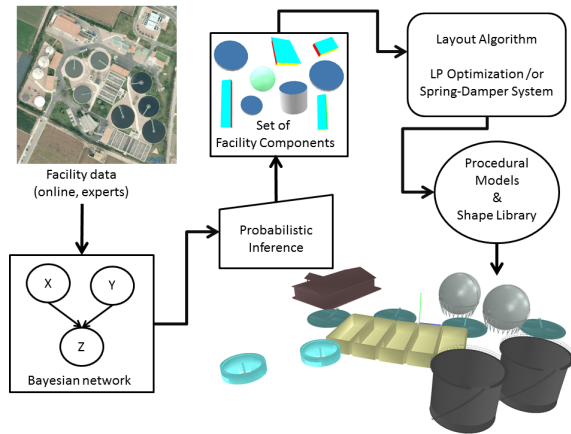


Figure 2: High level diagram of our system.

3.1. Data Collection

Python scripts were created to scrape websites for WWTP facility data (local population, flow rates, facility compo-

nents, etc.). The data were extracted from embedded HTML tags and stored in tables for use in constructing the BN. Table 1 shows a selected set of features and technical attributes of about 445 WWTP's located in Spain [Com13]. The set of feature attribute values for each WWTP are encoded using nominal (categorical) and numeric values.

3.2. Bayesian Network Representation

A BN provides the complete joint probability distribution for a set of random variables X_1, \dots, X_n . Each X_i is represented as a node in the BN with an associated conditional probability table (CPT) computed from the data set discussed in Section 3.1. The joint probabilities in a BN can be computed by multiplying together all the parent conditional probabilities $P(X_i | Parents(X_i))$ [RN10] taken from the CPT's:

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i | Parents(X_i)), \quad (1)$$

where n is the number of nodes in the BN and x_i is an attribute value of interest for each X_i in the network. We discretize continuous random variables prior to building the BN.

3.3. Learning the Bayesian Network

A BN network can be learned by iteratively modifying the network topology (adding edges, flipping edges, removing edges, etc.) and maximizing a network score (i.e. Bayesian information criteria) for each structure. While Merrell et al. [MSK10] use a learned BN for room adjacencies, we rely on the process flow diagram of a WWTP for connectivity and use structure learning only as an automation step to build the BN. The search over all possible networks is computationally expensive and in practice greedy algorithms are used to learn the network topology [Pe'05]. We use the K2 structure learning algorithm [CH92] as implemented in [HFH*09].

3.4. Inferring the Most Likely Components

A sampling of the BN conditioned on fixed attribute values allows for 3D model variation. For example, a user might want to create a facility for a city with a population of 1 million people. Given the conditioning, marginal probabilities $(P(X_1), \dots, P(X_n))$ can be computed from the joint distribution (Section 3.2) using the ancestor graph of X_i . The attribute value (x_i) with the maximum probability at each node in the BN is selected and placed in a set of components used for the layout discussed in Section 3.5. We use the Junction Tree Algorithm [LS98] as provided in [HFH*09] to compute marginal probabilities.

3.5. Arranging the Facility

We tested two methods to lay out the set of facility components from Section 3.4, a random physics based approach and a linear program (LP) mathematical optimization.

Feature Name	Nominal Attributes
Primary Treatment	{Decanting, Physical-Chemical}
Secondary Treatment	{Aerated Lagoons, Activated Sludge, Biocylinders, Biodiscs, Extended Aeration, ...}
Tertiary Treatment	{Infiltration-Percolation, Filtration, Floc-Coag, Reverse Osmosis, Ultrafiltration, ...}
Disinfection	{Chlorination, Ultraviolet}
Electricity Generation	{Cogeneration, Solar Panels}
Integer and Boolean Attributes	Feature Name
Integer types { \mathbb{Z} }	{Capacity, Population, Num. of Primary & Secondary Clarifiers, Num. Reactors, ...}
Boolean Types {Yes No}	{Degreaser, Decanting, Aerobic, Anaerobic, Centrifuge, Composting, Incineration, ...}

Table 1: Select features and attributes of the WWTP knowledge base.

3.5.1. Rigid Body Layout

The first method involves stochastically generating coordinates and assembling the components in a chain of spring constraints in a rigid body simulation [Bul13]. The resulting layout respects certain physical properties but does not optimize pipe costs, flow rates, or distances.

3.5.2. Linear Program Optimization

The second method uses LP optimization with linear constraints and a linear objective function, where the goal is to minimize $c^T x$ subject to $Ax \leq b$. The *decision variables* are the set of coordinates that minimize the Manhattan distance between components given the cost of piping (c), and flow (f) between the component parts (p). The system is constrained by a convex polygon site footprint and a set of separation distances (roads, paths, etc.). Each component is approximated with a bounding rectangle, or cell. We use the *BlockPlan LP* algorithm [MV88] as described by Heragu [Her08], with the objective function:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n c_{ij} f_{ij} (x_{ij}^+ + x_{ij}^- + y_{ij}^+ + y_{ij}^-) \quad (2)$$

where c_{ij} is the piping cost matrix and f_{ij} is the positive flow between components p_i and p_j . We assign a 0 flow between non-adjacent components and a random pipe cost [U.S00] between p_i and p_j because we do not have piping cost in our data set. To enforce non-negativity with respect to the distance [MV88], positive variables are introduced for both (x, y) and transformed to (x_{ij}^+, x_{ij}^-) and (y_{ij}^+, y_{ij}^-) . (See [Her08, MV88] for complete details.) Coordinates were mapped to world space and a procedurally generated shape or existing triangle-mesh placed at the centroid of each cell.

4. Results

Examples for three dynamically generated layouts are shown in Figures 3a - 3c. Figure 3d shows objects located at the center of cells whose coordinates were found using the LP optimization discussed in Section 3.5.2. The objects placed in the site are to simulate the actual components of a WWTP. The small short cylinders are primary clarifiers, the larger

ones are secondary clarifiers. The five rectangles represent the biologic reactors and the tallest cylinders are the digesters. The spheres are gas storage tanks and the long rectangle represents pre-treatment. The smallest rectangle symbolizes disinfection.

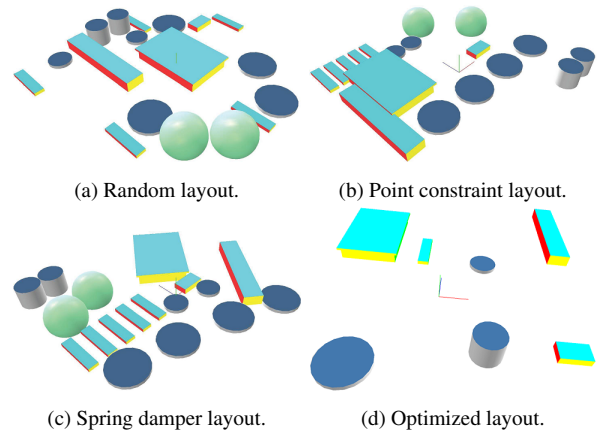


Figure 3: Various facility layouts

5. Conclusion & Future Work

We demonstrated the usefulness of arranging and modeling plausible WWTP's given a site and a probabilistic model. The most difficult aspects of this work were to identify a facility type, to construct a working knowledge base, and to find expert collaborators. Future work will include adding WWTP geometric details and textures to the data set and using a non-linear optimization with different distance metrics and boundary constraints to make the layouts more realistic.

6. Acknowledgements

This research was funded in part by a grant from the U.S. National Nuclear Security Administration's Non-proliferation University Collaboration and from two mobility stays funded by the Erasmus Mundus Programme of the European Commission's TEE Project. Special thanks to all the members of the Institute for Water and Environmental Engineering at UPV for their expertise in WWTP engineering.

References

- [BH93] BAUSBACHER E., HUNT R.: *Process plant layout and piping design*. PTR Prentice Hall, 1993. 1
- [BM99] BAILLARD C., MAÎTRE H.: 3-D reconstruction of urban scenes from aerial stereo imagery: A focusing strategy. *Computer Vision and Image Understanding* 76, 3 (1999), 244–258. 1
- [Bul13] Bullet physics engine. <http://bulletphysics.org/>, Oct 2013. Version 2.81. 3
- [Cen] CENTER FOR IMAGING SCIENCE, ROCHESTER INSTITUTE OF TECHNOLOGY: DIRSIG: Digital imaging and remote sensing image generation model. 1
- [CH92] COOPER G., HERSKOVITS E.: A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* 9, 4 (1992), 309–347. 2
- [CKGK11] CHAUDHURI S., KALOGERAKIS E., GUIBAS L., KOLTUN V.: Probabilistic reasoning for assembly-based 3D modeling. *SIGGRAPH* 30, 4 (2011), 35:1–35:10. 2
- [CL96] CURLESS B., LEVOY M.: A volumetric method for building complex models from range images. In *Proc. 23rd Conf. on CG and inter. tech.* (1996), ACM, pp. 303–312. 1
- [Com13] COMMUNITY OF VALENCIA, SPAIN: WWTP data. <http://www.epsar.gva.es>, May 2013. 2
- [FNSZ06] FORLANI G., NARDINOCCHI C., SCAIONI M., ZINGARETTI P.: Complete classification of raw LIDAR data and 3D reconstruction of buildings. *PAA Jour.* 8, 4 (2006), 357–374. 1
- [FRL*98] FAUGERAS O., ROBERT L., LAVEAU S., CSURKA G., ZELLER C., GAUCLIN C., ZOGHLAMI I.: 3-D reconstruction of urban scenes from image sequences. *Computer Vision and Image Understanding* 69, 3 (1998), 292–309. 1
- [FRS*12] FISHER M., RITCHIE D., SAVVA M., FUNKHOUSER T., HANRAHAN P.: Example-based synthesis of 3D object arrangements. *SIGGRAPH-ASIA* 31, 6 (Nov 2012), 135:1–11. 2
- [GDSS*11] GLEASON S., DEMA M., SARI-SARRAF H., CHERIYADAT A., VATSAVAI R., FERREL R.: Verification & validation of a semantic image tagging framework via generation of geospatial imagery ground truth. In *IGARSS* (2011), pp. 1577–1580. 1
- [GS09] GERMER T., SCHWARZ M.: Procedural arrangement of furniture for real-time walkthroughs. In *Computer Graphics Forum* (2009), vol. 28, pp. 2068–2078. 2
- [Her08] HERAGU S.: *Facilities Design*. CRC Press, 2008. 3
- [HFH*09] HALL M., FRANK E., HOLMES G., PFAHRINGER B., REUTEMANN P., WITTEN I. H.: The weka data mining software: An update. *SIGKDD Explorations* 11, 1 (2009). 2
- [HMG09] HAEGLER S., MÜLLER P., GOOL L. V.: Procedural modeling for digital cultural heritage. *Eurasip Journal on Image and Video Processing* 2009 (2009), 1–12. 2
- [HWA*10] HAEGLER S., WONKA P., ARISONA S. M., VAN GOOL L., MÜLLER P.: Grammar-based encoding of facades. In *Computer Graphics Forum* (2010), vol. 29, pp. 1479–1487. 1
- [KCKK12] KALOGERAKIS E., CHAUDHURI S., KOLLER D., KOLTUN V.: A probabilistic model for component-based shape synthesis. *SIGGRAPH* 31, 4 (Jul 2012), 55:1–55:11. 2
- [KW11] KELLY T., WONKA P.: Interactive architectural modeling with procedural extrusions. *ACM TOG* 30, 2 (2011), 14:1–14:15. 1
- [LS98] LAURITZEN S., SPIEGELHALTER D. J.: Local computations with probabilities on graphical structures and their application to expert systems. *J R Stat Soc* (1998), 157–224. 2
- [MSK10] MERRELL P., SCHKUFZA E., KOLTUN V.: Computer-generated residential building layouts. In *SIGGRAPH-ASIA* (New York, NY, USA, 2010), ACM, pp. 181:1–181:12. 2
- [MSL*11] MERRELL P., SCHKUFZA E., LI Z., AGRAWALA M., KOLTUN V.: Interactive furniture layout using interior design guidelines. *SIGGRAPH* (Aug 2011), 87:1–87:10. 2
- [MV88] MONTREUIL B., VENKATADRI U.: *From gross to net layouts: an efficient design model*. Quebec: Université Laval, Faculté des sciences de l'administration, 1988. 3
- [MWH*06] MÜLLER P., WONKA P., HAEGLER S., ULMER A., VAN GOOL L.: Procedural modeling of buildings. In *SIGGRAPH* (New York, NY, USA, 2006), ACM, pp. 614–623. 1
- [MZWVG07] MÜLLER P., ZENG G., WONKA P., VAN GOOL L.: Image-based procedural modeling of facades. *SIGGRAPH* 26, 3 (2007), 85. 1
- [NM11] NOMURA K., MIYATA K.: Automatic generation method for processing plants. *Forma* 26 (July 2011), 29–38. 2
- [Pe'05] PE'ER D.: Bayesian network analysis of signaling networks: a primer. *Sci STKE* 2005, 281 (2005), 14. 2
- [PKVG00] POLLEFEYS M., KOCH R., VERGAUWEN M., GOOL L. V.: Automated reconstruction of 3D scenes from sequences of images. *(ISPRS) Journal of Photogrammetry and Remote Sensing* 55, 4 (2000), 251–267. 1
- [PM01] PARISH Y. I. H., MÜLLER P.: Procedural modeling of cities. In *SIGGRAPH* (New York, NY, USA, 2001), ACM, pp. 301–308. 2
- [RN10] RUSSELL S., NORVIG P.: *Artificial Intelligence*, 3rd ed. Prentice Hall, 2010. 2
- [RPV*11] ROBERTS R., POPE P., VATSAVAI R., JIANG M., ARROWOOD L., TRUCANO T., GLEASON S., CHERIYADAT A., SOROKINE A., KATSAGGELOS A., PAPPAS T., GAINES L., CHILTON L.: Design of benchmark imagery for validating facility annotation algorithms. In *IGARSS* (2011). 1
- [RTP*10] ROBERTS R., TRUCANO T., POPE P., ARAGON C., JIANG M., WEI T., CHILTON L., BAKEL A.: On the verification and validation of geospatial image analysis algorithms. In *IGARSS* (2010), pp. 174–177. 1
- [TKS*13] TEBOUL O., KOKKINOS I., SIMON L., KOUTSOURAKIS P., PARAGIOS N.: Parsing facades with shape grammars and reinforcement learning. *IEEE Trans. on Pat. Anal. and Mach. Intel.* 35, 7 (2013), 1744–1756. 2
- [Tri13] TRIMBLE: Trimble 3D warehouse. <http://sketchup.google.com/3dwarehouse/>, 2013. 1
- [U.S00] U.S. EPA: Wastewater technology fact sheet: Pipe construction and materials, Sep 2000. EPA 832-F-00-068. 3
- [VAW*10] VANEGAS C. A., ALIAGA D. G., WONKA P., MÜLLER P., WADDELL P., WATSON B.: Modelling the appearance and behaviour of urban spaces. *Computer Graphics Forum* 29, 1 (2010), 25–42. 1
- [WMW*08] WATSON B., MULLER P., WONKA P., SEXTON C., VERYOVKA O., FULLER A.: Procedural urban modeling in practice. *IEEE CG & A* 28, 3 (2008), 18–26. 2
- [WOD09] WHITING E., OCHSENDORF J., DURAND F.: Procedural modeling of structurally-sound masonry buildings. *SIGGRAPH-ASIA* 28, 5 (2009), 112:1–112:9. 2
- [YYW*12] YEH Y.-T., YANG L., WATSON M., GOODMAN N. D., HANRAHAN P.: Synthesizing open worlds with constraints using locally annealed reversible jump MCMC. *SIGGRAPH* 31, 4 (2012). 2