

PENDETEKSIAN ANOMALY KEBIASAAN PELANGGAN MENGGUNAKAN SIGNATURE

Anggit Subekti¹, Moch Arif Bijaksana², Kiki Maulana³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Perkembangan telekomunikasi saat ini sangat pesat tetapi perusahaan terkadang mengalami kerugian karena adanya tindakan kejahatan. Salah satunya adalah fraud. Fraud dapat dideteksi dengan mengamati ketidaknormalan behaviour pelanggan atau anomali. Salah satu cara pendeteksian anomali dengan cara membuat pola data yang normal. Dalam kasus ini akan dibentuk pola kebiasaan pelanggan dengan menggunakan Call Detail Record (CDR) atau disebut signature.

Signature akan dijadikan acuan dalam pendeteksian anomali. Anomali terdeteksi jika data baru tidak sesuai dengan signature. Pembuatan signature dalam tugas akhir ini menggunakan Bayesian Network. Bayesian Network dapat mencari pola ketergantungan antar elemen signature. Bayesian Network ini terbukti dapat membuat pola signature yang dapat menggambarkan behaviour pelanggan dan mampu digunakan untuk mendeteksi anomali.

Kata Kunci : fraud , behaviour, anomali, CDR , Bayesian Network

Abstract

The growth of telecommunication technology in this time are increase progressively, but sometime company loss their money because of fraud. Fraud can be detected by inspecting an abnormal behaviour done by customer(anomaly). Anomaly can be detected by making a normal data pattern. In this study case the writer will create a pattern using Call Detail Record (CDR) or can be called as a signature.

Signature will be considerable as a parameter for anomaly detection. Anomaly will be detected if the pattern of a new data different from signature. In this Final Project signature is build using Bayesian Network. The writer use it because it can find the dependency of element of signature. It proved that Bayesian Network can made signature pattern which represent customer behaviour and can be use to detected an anomaly.

Keywords : fraud , behaviour, anomaly, CDR , Bayesian Network

Telkom
University

1. Pendahuluan

1.1 Latar belakang

Bisnis dalam dunia telekomunikasi saat ini sangat berkembang dengan cepat karena telekomunikasi sangat dibutuhkan oleh setiap orang. Namun, beberapa penelitian menyatakan bahwa perusahaan telekomunikasi banyak mengalami kerugian karena adanya *fraud*[11]. *Fraud* adalah sebuah tindak kejahatan untuk mendapat sebuah keuntungan dengan cara merugikan hak dan kepentingan orang lain. Perusahaan berusaha mendeteksi secepat mungkin beberapa kasus yang bisa menyebabkan *fraud*. *Fraud* dapat dideteksi dengan mengamati adanya ketidaknormalan pola penggunaan layanan telekomunikasi. Indikasi ketidaknormalan tersebut dapat muncul akibat keanehan pola penggunaan ataupun disebabkan oleh malfungsi sistem telekomunikasi itu sendiri[15].

Perusahaan membutuhkan biaya dan waktu yang banyak untuk mengidentifikasi *fraud* dikarenakan banyaknya data. Biasanya kasus *fraud* ini terjadi karena adanya penyimpangan *behaviour* atau kebiasaan pelanggan. Oleh karena itu permasalahan ini harus segera diatasi. Salah satu caranya dengan pendekatan data mining. Ada tiga permasalahan pokok yang harus dihadapi [11] antara lain:

1. Pendeteksian *fraud* biasanya dilakukan dengan menganalisa kebiasaan pelanggan. Kebiasaan pelanggan dapat dilihat pada *Call Detail Record*(CDR) yang disimpan pada perusahaan telekomunikasi. CDR merupakan kumpulan data pada saat seseorang sedang menelpon. Namun kendalanya jika menggunakan CDR tidak bisa menggambarkan kebiasaan setiap individu karena satu nomor telepon dapat digunakan oleh beberapa orang, misalnya telepon rumah. Pendeteksian *fraud* ini akan lebih akurat jika dilengkapi dengan data pelanggan, data pembayaran, dan lain-lain. Dengan data-data tersebut bisa membantu menganalisa kebiasaan pelanggan secara lengkap.
2. Data yang digunakan dalam pendeteksian adalah data yang *unsupervised* (data yang tidak terdapat kelas labelnya) maka kita harus dapat membedakan data yang normal dan yang *fraud*.
3. Kita membutuhkan *real time updates* data panggilan.

Salah satu cara untuk mendeteksi *fraud* dengan mendeteksi anomali pada kebiasaan pelanggan. Kebiasaan pelanggan bisa dijadikan *signature* yang menjadi identitas pelanggan tersebut. Pendeteksian anomali dilakukan dengan membandingkan kebiasaan pelanggan yang sekarang dengan *signature* yang telah terbentuk sebelumnya.

Oleh karena itu dalam tugas ini akan menganalisa pendeteksian anomali pada kebiasaan pelanggan dengan menggunakan *signature* dan data yang digunakan adalah CDR.

1.2 Perumusan masalah

Pokok permasalahan dalam tugas akhir ini adalah mendeteksi anomali pada kebiasaan pelanggan. Salah satu cara penyelesaian masalah ini adalah dengan membuat *signature* pelanggan. Salah satu cara untuk menggambarkan pola tersebut dengan *Bayesian Network*. *Bayesian Network* ini digunakan untuk mencari pola *signature* dengan cara menghitung probabilitas tiap komponen *signature* dan mencari hubungan antar komponen. Komponen *signature* merupakan data-data yang terdapat dalam CDR. Namun, sebelumnya data CDR akan dikonversi ke dalam pola-pola tertentu untuk mempermudah dalam pembentukan *network*. Data yang sudah terpolakan tersebut yang akan disimpan dalam data *history* sebagai *signature*-nya. Misalkan komponen *signature* adalah $X_1, X_2, X_3, \dots, X_n$ dan dengan menggunakan *Bayesian Network* diperoleh pola sebagai berikut: $P(X) = P(X_1) P(X_2 | X_1) \dots P(X_n | X_1, \dots, X_{n-1})$. Pola itulah yang digunakan sebagai *signature*. Pendeteksian anomali dilakukan dengan membandingkan CDR baru dengan pola yang telah terbentuk.

Berdasarkan permasalahan diatas maka terdapat rumusan masalah sebagai berikut:

1. Bagaimana cara menentukan *signature* pelanggan dengan menggunakan data *history*.
2. Bagaimana cara peng-*update*-an data *history* jika terdapat CDR baru.
3. Bagaimana cara pendeteksian anomali.

Adapun batasan masalah dalam pembuatan Tugas akhir ini antara lain:

1. Data yang digunakan dalam pembuatan *signature* adalah CDR yang telah dikonversi ke dalam pola-pola tertentu.
2. Sistem tidak dibuat untuk *real time*.
3. *signature* yang dibentuk hanya digunakan untuk pendeteksian anomali.
4. Penelitian tidak sampai pada pendeteksian *fraud*. Penelitian hanya pada pendeteksian anomali.

1.3 Tujuan

Tujuan dari pembuatan tugas akhir ini antara lain:

1. Membentuk *signature* dengan menggunakan *Bayesian Network*.
2. Menentukan cara peng-*update*-an data *history*.
3. Mendeteksi adanya kebiasaan pelanggan yang anomali dan tidak anomali .

1.4 Metodologi penyelesaian masalah

1. Studi Literatur

Tahapan studi literatur meliputi tahapan pengumpulan literatur yang mendukung penelitian ini sehingga di dapatkan dasar teori dan algoritma yang jelas dan tepat untuk penyelesaian masalah pembuatan *signature*, peng-*update*-an *signature* dan pendeteksian anomali kebiasaan pelanggan . Selain itu dalam tahapan ini dilakukan pengumpulan data untuk mendukung proses penelitian ini.

2. Perancangan Sistem

Tahapan ini digunakan untuk perancangan program untuk mengetahui anomali dengan menggunakan metode *signature*. Tahapannya meliputi pembuatan model sistem anomali *detection* dengan menggunakan metode *signature*. Hasil dari tahapan ini merupakan *blue print* yang akan digunakan untuk tahapan selanjutnya.

3. Implementasi Sistem

Tahapan implementasi digunakan untuk menuangkan *blue print* yang dibuat pada tahapan perancangan ke dalam program. Hasil dari tahapan ini berupa program yang dapat digunakan untuk mendeteksi anomali kebiasaan pelanggan.

4. Pengujian

Pengujian merupakan tahapan yang digunakan untuk melakukan pengujian pada sistem yang telah dibangun. Hasil dari tahapan ini berupa terdeteksinya pelanggan yang melakukan kebiasaan yang tidak normal.

5. Analisis

Analisis merupakan tahapan yang digunakan untuk melakukan analisa *signature* yang terbentuk apakah sudah mencerminkan pola perilaku pelanggan atau belum. Selain itu dianalisa juga setelah data diupdate *signature* masih mencerminkan pola perilaku yang sebelumnya dan pendeteksian anomali kebiasaan pelanggan.

6. Penyusunan Laporan

Hasil penelitian akan di susun menjadi suatu laporan yang meliputi aspek-aspek dalam penelitian yaitu teori dan implementasinya

5. Kesimpulan Dan Saran

5.1 Kesimpulan

1. *Bayesian Network* mampu membentuk pola *signature* yang dapat menggambarkan seluruh isi data.
2. *Bayesian Network* dapat mendeteksi anomali pada suatu data walaupun data hanya mengalami perubahan satu data saja.
3. Penambahan data histori dengan menggunakan isi data yang mempunyai pola yang sama dengan data histori tidak mempengaruhi pola *signature*.
4. Proses pembuatan *signature* membutuhkan waktu agak lama karena pada pembuatan jaringan semua variabel dibandingkan.

5.2 Saran

1. Sebaiknya *signature* dibentuk dengan menggunakan data yang disimpan dalam rentang waktu yang cukup lama sehingga *signature* dapat benar-benar menggambarkan kebiasaan pelanggan.
2. Sebaiknya *signature* dibentuk dengan menggunakan data CDR asli sehingga pendeteksian anomali dapat lebih akurat.
3. Pendeteksian anomali pada kebiasaan pelanggan sebaiknya tidak hanya menggunakan *inference bayes* saja karena *inference bayes* dapat mendeteksi satu perbedaan sehingga di khawatirkan banyak data yang terdeteksi sebagai anomali.
4. Sebaiknya sistem mampu melakukan peng-updatean secara *real time*.

Daftar Pustaka

- [1] Cahill M, Lambert D, Pinheiro J dan Sun D. 2002. *Detecting Fraud In The Real World*. dalam The Handbook of Massive Datasets. Kluwer Academic Publishers.pp911-930
- [2] Diana L, Pinheiro.J. *Mining A Stream of Transactions for Customer Patterns* . Bell Labs, Lucent Technologies.
- [3] Cliftonb, Chris . *Fraud and Misuse Detection*.12/22/06. www.cs.purdue.edu/homes/clifton/cs490d/Fraud.ppt
- [4] Data Mining Center.2006.*Aplikasi Data Mining Pada Telekomunikasi*.Bandung: Sekolah Tinggi Teknologi Telekomunikasi.
- [5] Douglas C.M , William W.1990. *Probabilitas Dan statistika Dalm Ilmu Rekayasa Mnajemen*.Penerbit Universitas Indonesia : Jakarta
- [6] Holm'en , Jaakko.2000.*User profiling and classification for fraud detection in mobile communications networks*. Helsinki : Helsinki University of Technology Department of Computer Science and EngineeringLaboratory of Computer and Information Science.
- [7] Kusuma dewi, Sri.2003.*Artificial Intelegence*.Yogyakarta: Grah Ilmu.
- [8] Osunmakinde I O.2005.*Telecommunications Fraud Detection usingBayesian Networks*. African Institute for Mathematical Sciences :South Africa.
- [9] Pedro Ferreira, Ronnie Alves, Orlando Belo and Luis Cortes~ao.2006.*Establishing Fraud Detection Patterns Based on signatures*. University of Minho : Department of Informatics, Campus of Gualtar 4710-057 Braga, Portugal
- [10] Saharon Rosset, Uzi Murad, Einat Neumann, Yizhak Idan, Gadi Pinkas.1999.*Discovery of Fraud Rules for Telecommunications - Challenges and Solutions*. KDD 99
- [11] Tan, Steinbach dan kumar.2004.*Data Mining Anomaly Detection*. Slide dari buku Introduction to Data Mining dalam kuliah data mining.
- [12] Richard E. Neapolitan.*Learning Bayesian Networks*.Northeastern Illinois University :Chicago
- [13] A.Suhendar,Hariman G. 2002. *Visual Modeling Menggunakan UML dan Rational Rose*.Penerbit Informatika Bandung
- [14] Wikipwdia.2007. Bayesian Network. www.wikipedia.com
- [15] Xerandy , Hendrawan.2006. *Telecommunication Fraud Di Indonesia : Kajian Masalah Dan Kebijakan Penanganannya*. Program Studi Teknik Elektro Sekolah Teknik Elektro dan Informatika : Institut Teknologi Bandung