Migration Letters

Volume: 20, No: S13(2023), pp. 322-329

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)
www.migrationletters.com

Scrutinize Analysis of Crime Investigation – Assigned Weighted Approach

Thiagarajan Kittappa[1*], S. Mrithul Snehal[2], J. Amrutha[3], J. John Osborne[4], A. Nishanth[5] and Anish Anburaja[6]

Abstract

This study focuses on trust rating between police records and observed current case's data, employing a weighted graph approach to make informed decisions in different criminal investigations. Four distinct cases, ranging from theft to homicide, were examined, each with its own assigned trust levels. The outcomes of these cases varied, with trust ratings playing a significant role in influencing the final decisions made by law enforcement agencies. The results underscore the importance of considering specific evidence and circumstances in the determination of suspects guilt or innocence in AI based outputs

Keywords: *Agent, Clients, Weighted Signed Graph.*

Introduction

In today's world, there are a lot of criminal offenses happening around the world. And these criminal offense data are useful for investigating the crime, and the availability ofdata pertaining to these criminal incidents plays a pivotal role in the process of crime investigation and the apprehension of offenders. By correlate the similarities between the data in the police record and the data collected by the crime scene investigators to determine if there was an AI based pattern between these criminal acts. If there is data available that supports this hypothesis. Department of Investigators can conclude that these acts were done by a proposed criminal entity.

The objective of the research is to enhance the operational efficiency and facilitate more accurate identification of potential suspects. This proposed algorithm discusses the trust level rating through aweighted sign approach to improve efficiency. The outcomes of this study hold significant promise for law enforcement agencies, enabling them to expedite and streamline the investigation process, ultimately contributing to the pursuit of justice in optimized manner.

Definitions

Simple Match Coefficient (SMC)

The term "simple match coefficient" is often used in the context of data matching and record linkage.It's a measure of how well two data records match or correspond. The simple match

[1*] Professor, Department of Mathematics, Rajalakshmi Institute of Technology, Chennai, India. vidhyamannan@yahoo.com
[2] Department of Artificial Intelligence and Data Science, Rajalakshmi Institute of Technology, Chennai, India. mrithulsnehal@gmail.com
[3] Department of Artificial Intelligence and Data Science, Rajalakshmi Institute of Technology, Chennai, India. amruthajaganathanjaganthan@gmail.com
[4] Department of Artificial Intelligence and Data Science, Rajalakshmi Institute of Technology, Chennai, India. osbornej406@gmail.com
[5] Department of Artificial Intelligence and Data Science, Rajalakshmi Institute of Technology, Chennai, India. nishanthtamil72@gmail.com
[6] Department of Information Sciences, University of Maryland, USA. anishanbu@gmail.com

coefficient is typically a binary measure, indicating whether two records match (1) or do not match (0) based on certain criteria or attributes. It's a straightforward way to assess similarity oragreement between data records.

Rao's Coefficient (RC)

Rao's coefficient, also known as Rao's quadratic entropy, is a statistical measure used to assess the dissimilarity or distance between two probability distributions. It is named after the Indian-Americanstatistician C. Radha krishna Rao.

Rao's coefficient is a mathematical way to quantify how two probability distributions differ fromeach other. It's often used in fields like statistics, information theory, and ecology to compare thesimilarity or dissimilarity between datasets. The formula for Rao's coefficient involves the use ofprobabilities from both distributions, and the result can range from 0 (indicating identical distributions) to a positive value (indicating increasing dissimilarity).

Jaccard Coefficient (JC)

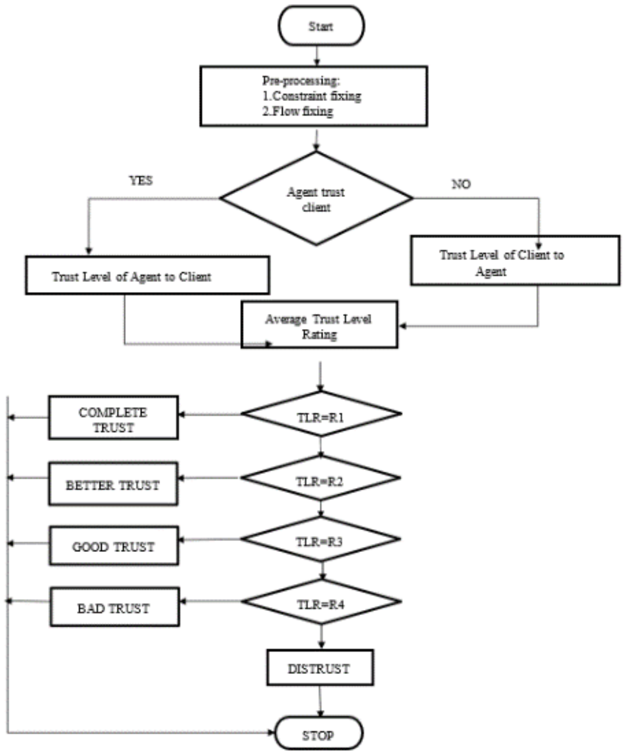
The Jaccard coefficient produces a value between 0 and 1, where 0 indicates no similarity (the sets have no elements in common), and 1 indicates complete similarity (the sets are identical). It's a usefulmetric for comparing the similarity of data samples or sets in various applications, such as text analysis, recommendation systems, and biology.

- **SMC**: S (i , j) = (P+S)/(P+Q+R+S)
- **RC:** S (i,j) = P/(P+Q+R+S)
- **JC**: S (i j) = P/P+Q+R)

Table 1.1: Consider some examples for table 1.1

PA	DESCRIPTION
++	Past Record Matches with Present
+ -	Past Record Matches but The Newly Collected Data Has Some Percentage of Variations.
- +	First Time the Record Is Being
- -	Past Record Doesn't Match with Present Collected

Flow Diagram of Proposed System



Examples

Case (I)

Necessary records for this case,

1) Past Record (PR)
2) Newly Collected Record (NCR)

Type of Crime is Theft

Assume that, the police database has case records on various types of crimes which are committed by an individualor a group of people. In this scenario consider theft to be the crime.

PR	NCR	TL	TR	DECISON
++	++	++	**P1**	**COMPLETE TRUST LEVEL**
++	+-	+-	**P2**	
++	-+	-+	**P3**	
++	--	--	**P4**	
(P1+P2+P3+P4)/4 = COMPLETE TRUST LEVEL				

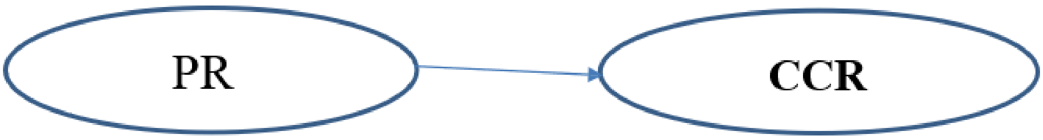
Use cases

- (++); The police concluded in the report that the crime happened between 11pm-1am and there is data available about the crimes that happened between 11pm-1am in database and the data's whichareavailable in database conveys that theft is the major cause committed between 11pm- 1am
- (+-); The police also reported that the crime happened in a nearby shop(specific). Even though the databasehas complete data about the crime happened in that area but the data about shop (specific) isnot present in the database so the similarity decreases.
- (-+); The police reported that these are foot prints along the pathway and collected the data but in thepastrecord there is no data which has the footprints of the criminals.
- (--); An outsider was found roaming in the area around the above-mentioned time but his shoe sizeislargerthan the footprints found at the crime scene and the database has data about the shoe size ofcriminals but not above-mentioned size.

Cases (II)

Necessary records for this case,

1) Past Record (PR)
2) Current Case Record (CCR)



Type of Crime is Murder

Consider an example in which crime scenario is taken has murder. Assume that the police of particular Jurisdiction suspect a person for murder. The weapon used for murder was a knife. This knife is similar to thatof knife which is used in meat shop and the person who is accused for the murder is a meat shop owner.

PR	CCR	TL	TR	DECISON
+-	++	+-	**P2**	**GOOD TRUSTLEVEL**
+-	+-	-+	**P3**	
+-	-+	--	**P5**	
+-	--	--	**P5**	
(P2+P3+P5+P5)/4 = GOOD TRUST LEVEL				

Use Cases

- If the client data matches with the agent but the person is deceptively accused, so the trust levelwill be (+-)
- If the client data have a similarity to the victim and the police believe that the person is the onewho committed the crime then the trust level will be (-+).
- If the client data doesn't have a similarity to the agent but the police accusses the person to be thevictim then the trust level will be (--)
- If the client data data doesn't match with the person but the police believes that the person was thevictim (--)

Case (III)

Type of Crime is Kidnapping

Assume, that there is a kidnapping case at some specific place which is being handled by police of the same Jurisdiction. They believed that the kidnapping case is related to Maruti Omni van. Hence the police suspected the driver as the criminal but he was innocent.

PR	CCR	TL	TR	DECISON
-+	++	-+	**P3**	**BAD TRUSTLEVEL**
-+	+-	--	**P5**	
-+	-+	--	**P5**	
-+	--	--	**P5**	
(P3+3P5)/4 = BAD TRUST LEVEL				

Use Cases

- If the client data matches even though the police does not suspect the criminal, then the trust levelwill be (-+)
- If the client data have similarities with the criminal but not with the person who was accused bythe police to be a criminal then the trust level will be (--)
- If the client data have similarities with the criminal accused by the police then the trust level will be (--) • If the client data doesn't have similarities with the criminal then the trust level will be (--).

Case (IV)

Necessary records for this case,

1) Police Doctrine (PD)
2) Behavior Of Suspection (BS)

Type of Crime is Homicide

Consider, that the police had collected the evidence of a murderer who engaged in the unlawful killing, but the evidence which they extracted has no reliability to the suspected personality, but there is either or possibleoffender who deliberate the suspection with the previous or present actions.

PD	BS	TL	TR	DECISON
--	++	-+	P3	BAD TRUSTLEVEL
--	+-	--	P5	
--	-+	--	P5	
--	--	--	P5	
P3+3P5)/4 = BAD TRUST LEVEL				

Use Cases

- The murder has been done by the client, but the police doesn't suspect him, the trust level willbe (-+).
- The murder has done in previous cases but not at present, yet the police doesn't suspect him,the trust level will be (--).
- The murder has not done in previous cases but at present it had been done, yet the policedoesn't suspect him, the trust level will be (--).
- The murder has not been done by the client, but the police doesn't suspect him, so the trustlevel will be (--).

Respective variablecase	Data in police record	Collected currentcase data
P	+	+
Q	+	-
R	-	+
S	-	-

Experimental Results and Comparison with Existed Records

For Case (I)

Observed data;

P	Q	R	S
100	75	50	25
SMC	RC	JC	
1/2	2/5	4/9	
CONCLUSION: BADTRUST LEVEL			

For Case (II)

P	Q	R	S
100	75	50	0
SMC	RC	JC	
0	0	0	
CONCLUSION: BADTRUST LEVEL			

For Case (III)

P	Q	R	S
0	0	25	0
SMC	RC	JC	
0	0	0	
CONCLUSION: BAD TRUST LEVEL			

For Case (IV)

P	Q	R	S
0	0	25	0
SMC	RC	JC	
0	0	0	
CONCLUSION: BAD TRUST LEVEL			

Conclusion

Case I

The final decision may be influenced by the trust ratings of the agents involved. In this case, person 1complete trust and the high trust level of person 2 would hold more weight, making the conclusion about the time of the crime and the absence of shop-specific data relatively more reliable. However, conclusions related to footprints and the outsider's shoe size may beless reliable due to the lower trust levels. The average trust rating would depend on how muchweight is assigned to each agent's input.

Case II

In the murder case, the trust levels vary depending on the situation, with the highest trust level (- +) indicating a strong belief in the accused person's guilt and the lowest trust level

(- -) suggesting a lack of evidence and similarity between the accused and the victim. The final decision would depend on the specific evidence and circumstances surrounding the murder case.

Case III

In the kidnapping case, trust levels vary, but they tend to be low, due to the possibility of wrongful accusations and the need for further investigation to determine the actual perpetrator. The final decision should consider concrete evidence and facts in order to establish the innocence or guilt of the accused driver.

Case IV

In this homicide case, trust levels vary, but they generally tend to be low, as the police donot suspect the client because the collected evidence is not reliable to the suspected personality. The final decision should consider the existing evidence in the given scenario andfacts to determine the true culprit, to analyze whether the suspected person is involved ornot.

Nomenclature

PA: Precedence of Acceptance

PR: Past Record

NCR: Newly Collected Record

CCR: Current Case Record

PD: Police Doctrine

BS: Behavior of Suspection

SMC: Simple Match Coefficient

RC: Rao's Coefficient

JC: Jaccard's Coefficient

TR: Trust Rating

TL: Trust Level

Acknowledgement

The authors would like to thank Prof. Ponnammal Natarajan worked as Director – Research, Anna University-

Chennai and Dr. Sundar R Currently Director, Rajalakshmi Institute of Technology, Chennai, for their intuitive ideas and fruitful discussions with respect to the paper's contribution.

References

Mohammad Reza Keyvanpour, Mostafa Javideh, Mohammad Reza Ebrahimi "Detecting and investigating crime by means of data mining: a general crime matchingframework". Department of Computer Engineering, Alzahra University, Tehrun, Iran Islamic Azad University, Qazvin Branch, Qazvin Iran.

K. Thiagarajan, A. Ragunathan, Ponnammal Natarajan, G. Poonkuzhali and Prashant Ranjan"

Weighted Graph Approach for Trust Reputation Management", World academy of science, Engineering and Technology 56 2009.
G.C. Oatley, J. Zeleznikow, B.W. Ewart, "Matching and Predicting Crimes," In Applications and Innovationsin Intelligent Systems XII in Proceedings of AI2004, The Twenty-fourth SGAI International Conference on Knowledge Based Systems and Applications of Artificial Intelligence. Ann Macintosh, Richard Ellis and Tony Allen Ed. London: Springer, pp. 19-32, 2004.
R. William Adderley, "The use of data mining techniques in crime trend analysis and offender profiling," Ph.D. thesis, University of Wolverhampton, Wolverhampton, England, 2007.
Y. Xiang, M. Chau, H. Atabakhsh, H. Chen, "Visualizing criminal relationships: comparison of a hyperbolic tree and a hierarchical list," Decision support systems,Elsevier Science Publishers, vol. 41 no.1, pp: 69-83, Nov. 2005.
Chen, W. Chung, Y. Qin, M. Chau, J. Xu, G. Wang, R. Zheng, and H. Atabakhsh, "Crime Data Mining: An Overview and Case Studies," in Proceedings of the 3rd NationalConference for Digital Government Research (dg.o 2003), pp. 1-5, Boston,MA, May 18-21, 2003.
H. Chen, H. Atabakhsh, T. Petersen, J. Schroeder, T. Buetow, L. Chaboya, C. O'Toole, M. Chau, T. Cushna, D. Casey, Z. Huang, "COPLINK: Visualization for Crime Analysis," ACM International Conference proceeding Series, Proceedings of the 2003 annual national conference on digital government research, Vol. 130, pp 1-6, Boston, MA, 2003.
R.V. Hauck, H. Atabakhsh, P. Ongvasith, H. Gupta, H. Chen, "Using Coplink to Analyze Criminal-Justice Data," Computer, vol. 35, no. 3, pp. 30-37, Mar. 2002.
H. Chen, W. Chung, J.J Xu, G. Wang, Y. Qin and M. Chau, "Crime Data Mining: A General Framework and Some Examples," Computer, vol. 37, no. 4, pp. 50-56, Apr. 2004.
J. Han and M. Kamber, Data Mining Concepts and Techniques, 2nd ed. Morgan Kaufmann, Nov. 3, 2005.
D. J. Hand, H. Mannila, P. Smyth, Principles of Data Mining, Massachusetts, MIT Press, 2001.
G.K. Gupta, Introduction to Data Mining with Case Studies, New Delhi: Prentice-hall of India, 2006.
Y.J. Oyang, C.Y. Chen, T.W. Yang, "A Study on the Hierarchical Data Clustering Algorithm Based on Gravity Theory," Proceedings of the 5th European Conference on Principles of Data Mining and Knowledge Discovery, p.350-361, Sep. 03-05, 2001.
J. Huysmans, D. Martens, B. Baesens, J. Vanthienen and T. Van Gestel, "Country Corruption Analysis with Self-Organizing Maps and Support Vector Machines," Proceedings of the Tenth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2006), Workshop on Intelligence and Security Informatics (WISI), vol. 3917, pp. 103-114, Singapore: Springer-Verlag, 2006.
M. Aurélio S. da Silva, A. M. V. Monteiro and J. S. Medeiros, "Visualization ofGeospatial data by component planes and U-Matrix," 6th Brazilian Symposium onGeoinformatics, São Paulo, Brazil, 22-24 Nov. 2004.
R. Neumayer, R. Mayer, G. Pölzlbauer and A. Rauber, "The Metro Visualisation of Component Planes for Self-Organising Maps," In Proceedings of the 20th International Joint Conference on Neural Networks (IJCNN'07), Orlando, FL, USA: IEEE Computer Society, Aug.12– 17, 2007, pp. 2788-2793.
M.R. Keyvanpouret al. / Procedia Computer Science 3(2011) 872–880 879 Mohammad Reza Keyvanpour / Procedia Computer Science 00 (2010) 000–000
G. Pölzlbauer, M. Dittenbach, A. Rauber, "Gradient visualization of grouped component planes on the SOM lattice," In Proceedings of the 5th Workshop on Self-Organizing Maps Paris (WSOM 2005), Sep. 5-8 2005, Paris, France, pp. 331-338.
K. l. Priddy, p. e. Keller, Artificial Neural Networks, New Delhi: Prentice-hall of India, 2007.
K. Hornik," Approximation capabilities of multi-layer feedforward networks," Neural Networks, vol. 4, no.2, pp. 251–257. Oxford, UK, Elsevier Science Ltd. 1991.
Z. Tang and J. MacLennan, Data Mining with SQL Server 2005, Indianapolis, USA: Wiley Publishing, 2005.
Bin Yu, Munindar P.Singh, An Evidential model for distributed reputation Management, AAMAS'02, July 15-19,2002, Bologna, Italy.
B. Yu and M. P. Singh. Trust and reputation management in a small- world network. In Proceedings of the 4th International Conference on MultiAgent Systems, pages 449450, 2000.
B.Yu and M.P.Singh. A social Mechanism of reputation management in electronic communities. In proceedings of the 4th International Workshop on Cooperative Information Agents, 2000.
E.Sampathkumar, (Eds: M.M Shikare and B.N.Waphare) "Semigraphs" Narosa Publishing house, New Delhi, (2004), 125-138.

F. Harary Addision-Wesley, Reading "Graph Theory", Mass, (1969).
G. Zacharia and P. Maes. Trust management through reputation mechanisms. Applied Artificial Intelligence, 14:881-908, 2000.
K. Aberer and Z. Despotovic. Managing trust in a peer-2-peer information system. In Proceedings of the10th International Conference on Information and Knowledge Management (CIKM), pages 310-317, 2001.
L. Xiong and L.Liu. Vulnerabilities and security issues in structured overlay networks: A quantitative analysis. In proceedings of the Annual Computer Security Applications Conference (ACSAC), 2004.
Rajesh Turlapati; Huhns, M.N., Multiagent reputation management to achieve robust software using redundancy. Intelligent Agent Technology, IEEE/WIC/ACM International Conference on Volume, Issue, 19- 22 Sept. 2005 Page(s): 386-392
Zhengqiang Liang and Weisong Shi, Enforcing Cooperative Resource Sharing in Untrusted P2P Computing Environments.876-880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar